







236  
27.8.70















# THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY  
BORIS SEMEONOFF

WITH THE ASSISTANCE OF  
C. BURT

AIDED IN THE SELECTION OF PAPERS BY

F. C. BARTLETT	G. C. DREW	T. H. PEAR
E. A. BOTT	H. J. EYSENCK	G. RYLE
D. E. BROADBENT	L. S. HEARNshaw	W. J. H. SPROTT
R. S. CREED	M. G. KENDALL	R. H. THOULESS
J. DREVER	O. A. OESER	M. D. VERNON

ASSISTANT EDITOR  
R. M. MCKENZIE

VOLUME 53, 1962

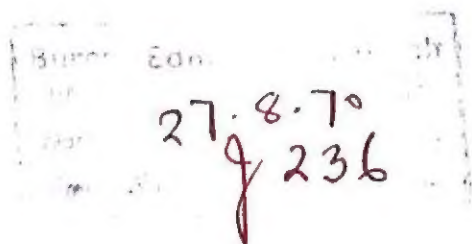


CAMBRIDGE  
AT THE UNIVERSITY PRESS  
1962



PUBLISHED BY  
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, N.W. 1  
American Branch: 32 East 57th Street, New York 22, N.Y.



*Printed in Great Britain at the University Press, Cambridge  
(Brooke Crutchley, University Printer)*

# CONTENTS OF VOLUME 53

## PART 1. FEBRUARY 1962

	PAGE
The perception of shape as a function of inclination. By R. B. JOYNSON and L. JOHN NEWSON . . . . .	1
The effects of meprobamate on kinesthetic figural after-effects. By C. G. COSTELLO . . . . .	17
Uncertainty and epistemic curiosity. By D. E. BERLYNE . . . . .	27
Experimentally induced drive and difficulty level in serial rote learning. By R. A. WILLETT and H. J. EYSENCK . . . . .	35
Tolerance for unrealistic experiences: a study of the generality of cognitive control. By GEORGE S. KLEIN, RILEY W. GARDNER and HERBERT J. SCHLESINGER . . . . .	41
The dynamic structure of attitudes in adults: a description of some established factors and their measurements by the motivational analysis test. By R. B. CATTELL, J. HORN and H. J. BUTCHER . . . . .	57
Skill and judgement of footballers in attempting to score goals. By JOHN COHEN and E. J. DEARNALEY . . . . .	71
Publications recently received . . . . .	89

## PART 2. MAY 1962

The absolute threshold of electric shock. By ROBERT T. GREEN . . . . .	107
Some investigations of perception of movement and related depth phenomena. By J. L. ZAJAC . . . . .	117
Control, defence and centration effect: a study of scanning behaviour. By RILEY W. GARDNER and ROBERT I. LONG . . . . .	129
Stimulus wavelength variation and size and distance judgements. By RAY OVER . . . . .	141
Disinhibition and the reminiscence effect in a motor learning task. By S. RACHMAN . . . . .	149
Is the serial-position curve invariant? By ARTHUR R. JENSEN . . . . .	159
Some properties of behaviour under fixed ratio and counting schedules. By H. M. B. HURWITZ . . . . .	167
An experimental study of the growth of some logical structures. By K. LOVELL, B. MITCHELL and I. R. EVERETT . . . . .	175
Effects of a subsidiary task on performance involving immediate memory by younger and older men. By D. E. BROADBENT and ALASTAIR HERON . . . . .	189
Publications recently received . . . . .	199



## PART 3. AUGUST 1962

	PAGE
Editorial. The <i>B.J.P.</i> in Jubilee Year . . . . .	221
Personalities in the early days of the British Psychological Society. By T. H. PEAR . . . . .	223
The concept of consciousness. By CYRIL BURT . . . . .	229
An exceptional talent for calculative thinking. By IAN M. L. HUNTER . . . . .	243
A study of religious belief. By L. B. BROWN . . . . .	259
An experimental study of Piaget's theory of the development of number in children. By H. BLAIR HOOD . . . . .	273
Group influence on the perception of ambiguous stimuli. By A. E. M. SEABORNE . . . . .	287
Conditioning and personality. By H. J. EYSENCK . . . . .	299
A theory of the serial position effect. By EDWARD A. FEIGENBAUM and HERBERT A. SIMON . . . . .	307
Perceptual-motor transfer in imbeciles: a second series of experiments. By A. D. B. CLARK and MARGARET COOKSON . . . . .	321
Changes in autokinetic perception as a function of the transfer of conditioning effects. By BOBBY J. FARROW and JOHN F. SANTOS . . . . .	331
Publications recently received . . . . .	339

## PART 4. NOVEMBER 1962

A two-factor theory of vigilance. By P. D. McCORMACK . . . . .	357
GSR Conditioning and pseudoconditioning. By IRENE MARTIN . . . . .	365
Relevance and category scales of judgement. By ROBERT S. DAVIDON . . . . .	373
Cognitive controls of attention and inhibition: a study of individual consistencies. By RILEY W. GARDNER and ROBERT I. LONG . . . . .	381
Sequential recall of a mixed list. By ROBERT T. GREEN and GRAHAM HARDING . . . . .	389
Transfer between difficult and easy tasks. By D. H. HOLDING . . . . .	397
Peripheral vision, refractoriness and eye movements in fast oral reading. By E. C. POULTON . . . . .	409
On the psychological nature of stage impersonation. By R. NATADZE . . . . .	421
Brightness judgements and stimulus size and distance. By RAY OVER . . . . .	431
The perception of rhythmically repeated linear motion in the horizontal plane. By E. G. WALSH . . . . .	439
Perceptual judgement as a function of context. By LUDWIG IMMERGLUCK . . . . .	447
The effect of reserpine on conditioned fear responses. By R. D. SAVAGE . . . . .	451
Correspondence. From H. J. EYSENCK, G. A. FOULDS and J. G. INGHAM . . . . .	455
Editorial Note . . . . .	460
Publications recently received . . . . .	461
Other publications received . . . . .	473

## INDEX OF AUTHORS

	PAGE		PAGE
BERLYNE, D. E. . . . .	27	HURWITZ, H. M. B. . . . .	167
BROADBENT, D. E. . . . .	189	IMMERGLUCK, L. . . . .	447
BROWN, L. B. . . . .	259	INGHAM, J. G. . . . .	458
BURT, C. . . . .	229	JENSEN, A. R. . . . .	159
BUTCHER, H. J. . . . .	57	JOYNSON, B. B. . . . .	1
CATTELL, R. B. . . . .	57	KLEIN, G. S. . . . .	41
CLARKE, A. D. B. . . . .	321	LONG, R. I. . . . .	129, 381
COHEN, J. . . . .	71	LOVELL, K. . . . .	175
COOKSON, M. . . . .	321	MCCORMACK, P. D. . . . .	357
COSTELLO, C. G. . . . .	17	MARTIN, I. . . . .	305
DAVIDON, R. S. . . . .	373	MITCHELL, B. . . . .	175
DEARNALEY, E. J. . . . .	71	NATADZE, R. . . . .	421
EVERETT, I. R. . . . .	175	NEWSON, L. J. . . . .	1
EYSENCK, H. J. . . . .	35, 299, 455	OVER, R. . . . .	141, 431
FARROW, J. . . . .	331	PEAR, T. H. . . . .	223
FEIGENBAUM, E. A. . . . .	307	POULTON, E. C. . . . .	409
FOULDS, G. A. . . . .	456	RACHMAN, S. . . . .	149
GARDNER, R. W. . . . .	41, 129, 381	SANTOS, J. F. . . . .	331
GREEN, R. T. . . . .	107, 389	SAVAGE, R. D. . . . .	451
HARDING, G. . . . .	389	SCHLESINGER, H. J. . . . .	41
HERON, A. . . . .	189	SEABORNE, A. E. M. . . . .	287
HOLDING, D. H. . . . .	397	SIMON, H. A. . . . .	307
HOOD, H. B. . . . .	273	WALSH, E. G. . . . .	439
HORN, J. . . . .	57	WILLETT, R. A. . . . .	35
HUNTER, I. M. L. . . . .	243	ZAJAC, J. L. . . . .	117

## INDEX OF PUBLICATIONS REVIEWED

- |   |   |
|---|---|
| <p>ADCOCK, C. J. <i>Fundamentals of Psychology</i> 207</p> <p>ADLER, G. (ed.). <i>Current Trends in Analytical Psychology</i> 94</p> <p>ADLER, G. <i>The Living Symbol</i> 209</p> <p>ANASTASI, A. <i>Psychological Testing</i> (2nd edition) 343</p> <p>ARMSTRONG, D. M. <i>Perception and the Physical World</i> 200</p> <p>ARNOLD, M. B. <i>Emotion and Personality</i> 342</p> <p>BACHRACH, A. J. (ed.). <i>Experimental Foundations of Clinical Psychology</i> 468</p> <p>BALDAMUS, W. <i>Efficiency and Effort</i> 102</p> <p>BANTON, M. (ed.). <i>Darwinism and The Study of Society</i> 464</p> <p>BELLAK, L. <i>Contemporary European Psychiatry</i> 217</p> <p>BELLER, E. K. <i>Clinical Process</i> 354</p> <p>BERNE, E. <i>Transactional Analysis in Psychotherapy</i> 217</p> <p>BERNHARDT, K. S. (ed.). <i>Training for Research in Psychology</i> 98</p> <p>BLUM, G. S. <i>A Model of the Mind</i> 345</p> | <p>BRAIN, Sir RUSSEL. <i>Speech Disorders</i> 343</p> <p><i>The British Journal of Social and Clinical Psychology</i>, vol. 1, part 1 356</p> <p>BROWN, N. O. <i>Life against Death</i> 93</p> <p>BROWN, W. P. <i>Conceptions of Perceptual Defence</i> 463</p> <p>BURNS, T. &amp; STALKER, G. M. <i>The Management of Innovation</i> 104</p> <p>BUROS, O. K. (ed.). <i>Tests in Print</i> 343</p> <p>BURTON, A. <i>Psychotherapy of the Psychoses</i> 214</p> <p>CAPLAN, G. <i>An Approach to Community Mental Health</i> 214</p> <p>CAPLAN, G. <i>Prevention of Mental Disorders in Children</i> 214</p> <p>COFER, C. N. (ed.). <i>Verbal Learning and Verbal Behavior</i> 207</p> <p>COUÉ, E. <i>Better and Better Every Day</i> 105</p> <p>COVILLE, W., COSTELLO, T. W. &amp; ROURKE, F. L. <i>Abnormal Psychology</i> 105</p> <p>DALTON, R. H. <i>Personality and Social Interaction</i> 341</p> <p>DEHAAN, R. F. &amp; HAVIGHURST, R. J. <i>Educating Gifted Children</i> (revised edition) 103</p> |
|---|---|



- DIAMOND, S. (ed.). *Culture in History. Essays in Honor of Paul Radin* 102
- DRY, A. M. *The Psychology of Jung* 352
- Ergonomics*, vol. 5, no. 1 354
- EYSENCK, H. J. (ed.). *Behaviour Therapy and the Neuroses* 211
- EYSENCK, H. J. (ed.). *Experiments in Personality*, vols. 1 and 2 89
- EYSENCK, H. J. (ed.). *Handbook of Abnormal Psychology* 345
- FLUCKIGER, F. A., TRIPP, C. A. & WEINBERG, G. H. *A Review of Experimental Research in Graphology* 1933-60 355
- FOSS, B. M. (ed.). *Determinants of Infant Behaviour* 466
- FREEHILL, M. F. *Gifted Children: their Psychology and Education* 219
- FREUD, E. L. (ed.). *Letters of Sigmund Freud—1873-1939* 350
- FURNEAUX, W. D. *The Chosen Few* 199
- GALIFRET, Y. (ed.). *Mechanisms of Colour Discrimination* 98
- GLIDENWELL, J. C. (ed.). *Parental Attitudes and Child Behaviour* 211
- GLOVER, E. *The Roots of Crime* 92
- GLUECK, S. and E. *Family Environment and Delinquency* 349
- GORDON, R. *Stereotypy of Imagery and Belief as an Ego Defence* 463
- GUNTUP, H. *Personality Structure and Human Interaction* 339
- HAMLYN, D. W. *Sensation and Perception* 200
- HARRIS, T. L. & SCHWAHN, W. E. *Selected Readings on The Learning Process* 98
- HARVEY, O. J., HUNT, D. E. & SCHRODER, H. M. *Conceptual Systems and Personality Organization* 340
- HERON, A. & CHOWN, S. M. *Ageing and the Semi-skilled* 349
- HONKAVAARA, S. *The Psychology of Expression* 339
- HUBER, J. T. *Report Writing in Psychology and Psychiatry* 97
- Journal of Psychiatric Research*, Vol. 1, no. 1, 220
- KAGAN, J. & LESSER, G. S. *Contemporary Issues in Thematic Apperceptive Methods* 218
- KAINZ, F. *Die Sprache der Tiere* 205
- KATZ, B. & LEWIS, R. T. *The Psychology of Abnormal Behavior* (second edition) 95
- KELLOGG, W. N. *Porpoises and Sonar* 205
- KING, R. A. *Readings for an Introduction to Psychology* 354
- KLEEMEIER, R. W. *Aging and Leisure* 91
- KOCH, S. (ed.). *Psychology: A Study of a Science*, Vol. 4, Study II 461
- LAING, R. D. *The Self and Others* 209
- LANGEN, D. *Anleitung zur Gestuften Aktivhypnose* 218
- LAWRIE, M. *Nature Hits Back* 105
- LEIGH, D. *The Historical Development of British Psychiatry*, Vol. 1, 18th and 19th centuries 352
- LEONT'EV, A. N. & ZAPOROZHETS, A. V. *Rehabilitation of Hand Function* 472
- LOVELL, K. *The Growth of Basic Mathematical and Scientific Concepts in Children* 202
- LUNDIN, R. W. *Personality an Experimental Approach* 465
- LURIA, A. R. *The Role of Speech in the Regulation of Normal and Abnormal Behaviour* 90
- MACKEITH, R. & SANDLER, J. *Psychosomatic Aspects of Paediatrics* 90
- MADIGAN, M. E. *Psychology—Principles and Applications* (3rd edition) 465
- MALIS, G. Y. *Research on the Etiology of Schizophrenia* 344
- MAYS, J. B. *On the Threshold of Delinquency* 470
- MEREDITH, G. M. & WONG, G. W. *A Table of Values for the Freeman-Tukey Square-root Transformation* 355
- MEREDITH, P. *Teach Yourself: Learning, Remembering and Knowing* 102
- MORGAN, C. T. *Introduction to Psychology* (2nd edition) 354
- MORRIS, D. *The Biology of Art* 353
- MUNN, N. L. *Introduction to Psychology* 465
- MÜTZE, K. (ed.). *Brockhaus—ABC der Optik* 100
- NEMIAH, J. C. *Foundations of Psychopathology* 95
- NOBLE, C. E. *Measurements of Association Value (a), Rated Associations (a'), and Scaled Meaningfulness (m')* 355
- ODLUM, D. *Journey through Adolescence* 105
- PASK, G. *An Approach to Cybernetics* 347
- PEAR, T. H. *The Moulding of Modern Man* 208
- PEVZNER, M. S. *Oligophrenia. Mental Deficiency in Children* 472
- Problems of Psychology*, Numbers 1 and 2 220
- Productivity*, Vol. 1 356
- RAND, H. A. *Graphology* 105
- REMITZ, U. *Professional Satisfaction among Swedish Bank Employees* 219
- RESTLE, F. *Psychology of Judgment and Choice* 202
- RODGER, R. S. *Statistical Reasoning in Psychology* 97
- ROGERS, C. R. *On Becoming a Person: A Therapist's View of Psychotherapy* 91
- ROMMETVEIT, R. *Selectivity, Intuition and Halo Effects in Social Perception* 100
- ROSENBLITH, W. A. (ed.). *Sensory Communication* 218
- ROTHKOPF, E. Z. & COKE, E. U. *Intralist Association Data for 99 words of the Kent-Rosanoff Word List* 355
- RUBIN, T. I. *Jordi* 105
- RUSSELL, C. & RUSSELL, W. M. S. *Human Behaviour: A New Approach* 208
- RUSSELL, W. R. & ESPIR, M. L. E. *Traumatic Aphasia* 206
- SANFORD, F. H. *Psychology—A Scientific Study of Man* 465
- SCHERKE, F. *Der Politische Charakter* 217
- SCHMIDT, G., STECK, H. & BADER, A. *Though this be Madness* 216
- SEBEOK, T. A. (ed.). *Style in Language* 353
- SHIELDS, R. W. *A Cure of Delinquents* 471
- SIDMAN, M. *Tactics of Scientific Research Evaluating Experimental Data in Psychology* 204

- SIDNEY, E. & BROWN, M. *The Skills of Interviewing* 471
- SNYDER, W. U. & SNYDER, B. J. *The Psychotherapy Relationship* 215
- SUPPES, P. & ATKINSON, R. C. *Markov Learning Models for Multiperson Interactions* 203
- SUTHERLAND, N. S. *Shape Discrimination by Animals* 204
- SYMONDS, P. M. with JENSEN, A. R. *From Adolescent to Adult* 348
- TAYLOR, F. K. *The Analysis of Therapeutic Groups* 96
- TERMAN, L. M. & MERRILL, M. A. *Stanford-Binet Intelligence Scale Third Revision (1960)—Form L-M* 472
- THORNDIKE, R. L. & HAGEN, E. *Measurement and Evaluation in Psychology and Education* (2nd edition) 343
- TIZARD, J. & GRAD, J. C. *The Mentally Handicapped and their Families* 96
- TOMAN, W. *Family Constellation* 469
- TRASLER, G. *The Explanation of Criminality* 469
- VERNON, M. D. *The Psychology of Perception* 462
- VIAUD, G. *Intelligence—Its Evolution and Forms* 105
- Vision Research—An International Journal*, Vol. 1, nos. 1-2 100
- WEINBERG, S. K. *Social Problems in our Time* 101
- WESTMAN, H. *The Springs of Creativity* 209
- WISEMAN, S. (ed.). *Examinations and English Education* 104
- WOLSTENHOLME, G. E. W. & O'CONNOR, M. (eds.). *The Nature of Sleep* 92
- WOOD, M. M. *Paths of Loneliness* 105
- WYBURN, G. M. *The Nervous System* 348
- YOUNG, P. T. *Motivation and Emotion* 342





## THE PERCEPTION OF SHAPE AS A FUNCTION OF INCLINATION

BY R. B. JOYNSON AND L. JOHN NEWSON

*Department of Psychology, University of Nottingham*

This is a descriptive study of the various judgments of shape which are spontaneously made when the inclination of an object is varied. The full range of inclinations is examined, using independent and unsophisticated subjects at each inclination, and recording subjects' descriptions.

The chief findings are: (1) Two main judgments are possible: judgments of real shape (R), which approximate to the true shape at all inclinations; and judgments contrasted with these (not-real or N), which show typically a compromise between true and retinal shape. (2) Some subjects (group RN) are spontaneously aware of both possibilities. They describe the N judgment as an attempt to equate on the basis of a direct, instantaneous, visual impression; whereas the R judgment involves 'allowing for the slope'. These subjects predominate at greater inclinations. (3) The remaining subjects (group RO) are spontaneously aware of the R possibility only. For these subjects, the R judgment seems to approximate closely to that quality of direct visual impression which for group RN characterizes the N judgment.

It is suggested that judgments of real shape involve a skill. Group RN are in an 'investigating' frame of mind, and offer an analysis of this skill. Group RO, in an 'unreflecting' frame of mind, take it for granted. Determining factors are also noted in the experimental situation and the instructions. It is suggested that previous studies fail to take adequate notice of these factors, and so give partial accounts.

### I. INTRODUCTION

The main aim of this study is to describe the perceptual phenomena which accompany variations in the inclination of an object to the line of sight. The retinal image undergoes a projective transformation. Does judgment remain in accordance with the real shape of the object ('constancy')? Does it correspond to the projective shape? Does it show a compromise between these two? Or are there further possibilities? As shown below, previous work does not answer these questions satisfactorily. It does not provide a complete survey, over the full range of inclinations, of the various perceptual judgments which are in fact possible. The present study attempts this.

In conducting the study, special attention has been given to the following points.

(1) To provide a complete survey, it is necessary to examine the full range of inclinations, from  $0^\circ$  tilt (where the object is frontal-parallel) to as near  $90^\circ$  tilt as is practicable. This study examines tilts from  $0^\circ$  to  $80^\circ$ . (Some experimenters measure tilt in the reverse direction: our  $0^\circ$  tilt becomes  $90^\circ$  tilt. We have adopted the system which seems more widely used, and, in referring to other work, we have converted measurements into our own system where necessary.)

(2) An adequate number of subjects must be employed. Limits are arbitrary, but it is unsatisfactory to use less than ten subjects in a field where individual differences are frequently encountered. Twenty subjects are examined at each inclination.

(3) Since experience at one inclination may affect later judgments at another inclination, it is preferable, as here, to use independent subjects at each inclination (unless of course the aim is to study such interaction).



(4) Subjects should not know what results have been found in previous experiments, or what views may be held by the experimenter. Studies which use the experimenter himself as a subject, or his psychological colleagues, are open to serious criticism (cf. Vernon, 1952, p. 223). We have used subjects who are unfamiliar with the literature, and have not previously taken part in psychological experiments.

(5) The apparatus must not impose artificial limitations on the judgments which the subject can make. It is possible, for instance, to present the subject with a series of comparison shapes which does not include the standard shape, thus eliminating the possibility of 'constancy' (Gottheil & Bitterman, 1951; for the analogous case in size constancy, see Lambercier, 1946; Joynson, 1949).

(6) Experimental instructions need particular care, since a variety of judgments are possible. Sometimes precise instructions are used. This raises many problems. First, such instructions may well limit the types of judgment obtained, and so give an incomplete picture. Secondly, some subjects may not be able to obey them closely (cf. Vernon, 1952, p. 222), and may make a different type of judgment. It will therefore be necessary to question subjects to ascertain exactly what they have done. Even if the subject's judgment approximates to that prescribed by the instructions, it is not legitimate to assume that it is exactly the same; and the instructions themselves may have influenced the subject's description of his judgment. Indeed, the strongest objection to precise instructions is that the experimenter assumes, before the experiment has been carried out, that he knows what judgments are possible; so there is a danger that his preliminary introspections, perhaps determined in part by his theoretical views, may appear in the results as objective findings. In order to avoid these pitfalls, we have used intentionally vague instructions, namely to judge when the two objects 'look the same'. Subjects respond in various ways to these instructions. But their judgments are relatively spontaneous, and they are asked when they have made them to describe them in their own words (cf. Joynson, 1958a).

### *Apparatus*

The standard object was an isosceles triangle of red paper (base 4.5 in. by height 2.5 in.). It was mounted centrally on a white card (base 7 in. by height 9.5 in.), which was attached to a camera tripod so that the angle of inclination could be varied in the vertical plane from 0° (upright: frontal-parallel) to 90° (horizontal). The inclinations used were: 0°, 25°, 50°, 70°, 80°. The standard was placed at a distance of nine feet from the subject, and was initially screened from his view. A chin rest ensured that viewing took place without head-movements.

The comparison objects were a series of twenty-six triangles made of thin paxolin sheet. All had the same base as the standard (4.5 in.), but their heights varied in steps of 0.1 in. from 0.5 in. (No. 1 comparison triangle) to 3.0 in. (No. 26 comparison triangle). No. 21 triangle was the same shape as the standard triangle. No. 26 triangle was 0.5 in. taller than the standard. No. 1 triangle corresponded to the retinal shape of the standard when tilted through 80°. Thus whatever the tilt of the standard, the comparison series extended to shapes taller than the real shape of the standard, and, except at 80°, to shapes shorter than the retinal shape of the standard.

The comparison triangles were laid out on a table in serial order immediately in front of the subject. He looked down at them from a height of about 2 ft., so that all were originally presented to him at approximately 0° to 10° tilt, with a vertical separation from the standard of about 80°. The experiment was conducted under artificial illumination in a room 20 × 12 ft. Fig. 1 shows the general arrangement of apparatus in the laboratory.

*Procedure*

The subject was seated at the table, with his chin on the rest, and told that he would shortly be shown a triangular shape. The following instructions were next read: 'Please choose from the series of triangles arranged in front of you, the one that looks most like the one you are going to see'. The screen was then raised to reveal the standard. The subject was told that there was no time limit, and that comments and questions would be welcome. He was permitted to handle the comparison shapes, if he wished. Binocular vision was used. When he had made his judgment, he was asked 'Please describe what you have done'.

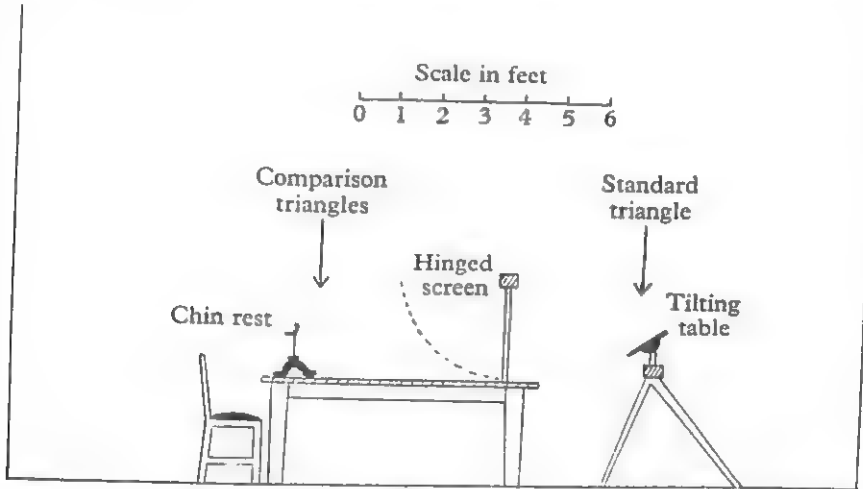


Fig. 1. Disposition of apparatus in laboratory.

One hundred subjects were examined, twenty at each of the five angles of inclination of the standard. The subjects were undergraduate or graduate members of this university. Two criteria were used in selecting them: first, ability to pass a standard test of visual acuity, with glasses if worn; secondly, absence of formal training in the psychology of perception, or previous experience of perceptual experiments.

It will be seen that the full range of tilts is examined; that twenty independent subjects are used at each inclination; that subjects have no special prior knowledge of psychology; that the apparatus permits the full range of possible judgments; and that subjects are not restricted by the instructions to any special type of judgment, but are free to describe what they have spontaneously done.

## II. QUANTITATIVE RESULTS

A variety of judgments were encountered. There were also considerable individual differences, varying with the degree of tilt, both in subjects' awareness of the different judgments and in their preference among them. In view of the complexity of the results, we shall first describe briefly the main types of judgment, and their quantitative value. We shall then consider how the results arose, the individual differences, and the more detailed description of the judgments.

Two main types of judgment were found. (1) The subject may select that com-

parison triangle which he considers to be the same 'real' shape as the standard. This will be termed a 'real' or R judgment. (2) The subject may make another judgment which is contrasted with the R judgment. This will be termed an 'other-than-real', 'not real', or N judgment. The phrase 'look the same shape' was used to refer to both judgments, and is therefore misleading if used without qualification. We shall return to this point. Here we note that the crucial question in classifying these judgments is whether the subject considers he is equating the objects for 'real' shape, or whether he considers that he is doing something other than this. In addition, N judgments may be made in two ways. Subjects may look back and forth from standard to variable, thus making a successive comparison (Nsc judgments). Alternatively, they may pick up the variables and compare them with the standard by placing them in almost the same line of sight, thus making a simultaneous comparison by juxtaposition (Nsm judgments). (R judgments were always made with successive comparison.)

Table 1. *Quantitative results*

Measure	Judgment	Angle of inclination									
		0°	<i>n</i>	25°	<i>n</i>	50°	<i>n</i>	70°	<i>n</i>	80°	<i>n</i>
Mean	R	22.40	20	21.60	20	22.05	20	21.90	20	20.87	20
	Nsc	—		17.67	3	18.00	10	15.47	15	12.09	16
	Nsm	—		—		13.00	1	6.33	3	7.00	2
Scatter ( $\sigma$ )	R	1.39		1.11		1.64		2.28		2.49	
	Nsc	—		2.02		1.84		1.94		3.56	
	Nsm	—		—		—		1.20		1.00	

The quantitative value of these judgments, at the different inclinations, is given in Table 1, and in Figs. 2 (R judgments) and 3 (N judgments). It will be seen that all subjects are able to make R judgments at every inclination, but N judgments follow a different pattern: there is no inclination at which all subjects can make N judgments, and at 0° no subject can make N judgments.

(1) *R judgments.* There is a small increase in scatter with inclination (Bartlett's test for homogeneity of variance yields  $\chi^2 = 12.4$ ;  $P < 0.02$ ). At 0°, the mean of the judgments overestimates the true height of the standard shape ( $t = 4.38$ ;  $P < 0.001$ ). This indicates a constant error, possibly due to the fact that the standard triangle was mounted on a contrasting white background. This constant error may be assumed to operate at every inclination. There is no evidence for any significant trend in mean values with inclination; indeed, the majority of R judgments at 80° fall within the range obtained at 0°. In short, R judgments show little change with inclination, and their close approximation to the 'real' shape, at every inclination, is notable.

(2) *Nsc judgments.* The data perhaps suggest an increase in scatter with inclination, but this is not significant. In mean value, these judgments are less than R judgments: there is no instance of an observer giving a smaller value for his R judgment than for his Nsc judgment. Nsc judgments show typically a compromise between the 'real' and the 'retinal' or 'projective' shape. In view of the considerable scatter which these judgments show, the question whether the degree of compromise varies with inclination has not been examined.



(3) *Nsm judgments*. Numbers are too small to reach any conclusions about scatter. In mean value, these judgments approximate more closely than Nsc judgments to the correct 'projective' shape.

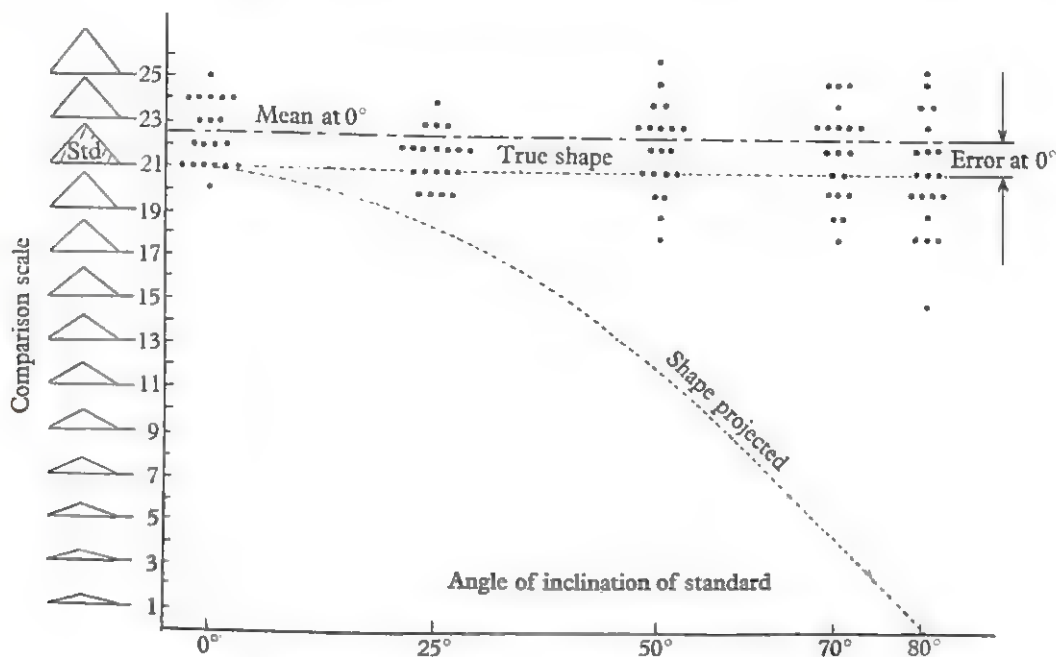


Fig. 2. Distribution of real (R) judgments.

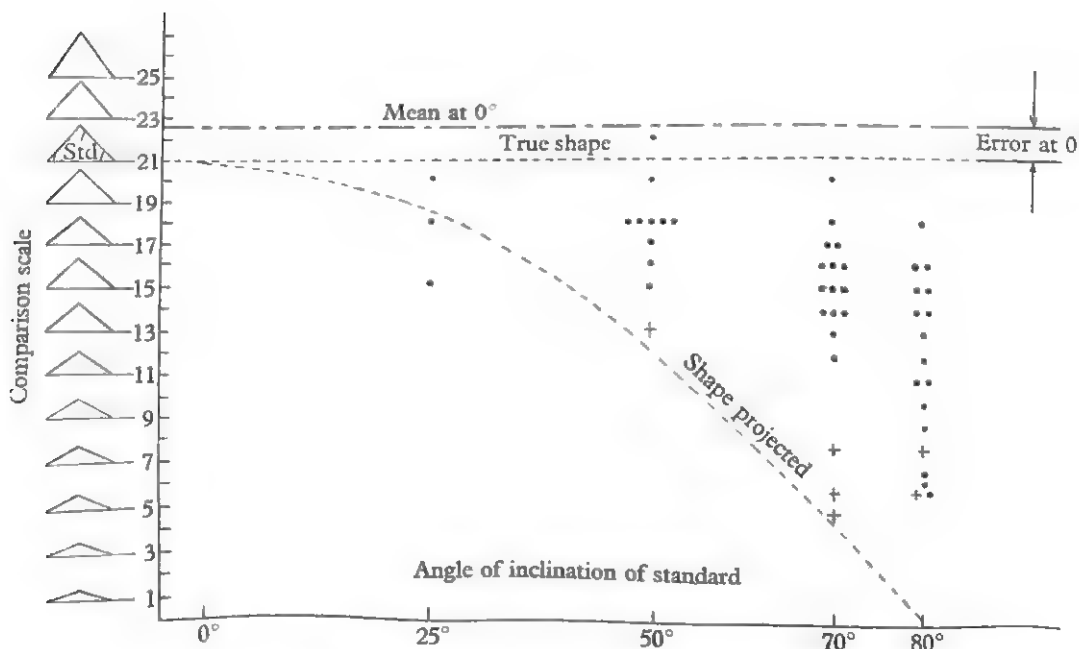


Fig. 3. Distribution of other-than-real (N) judgments.

●, Nsc judgments; +, Nsm judgments.

## III. INDIVIDUAL DIFFERENCES

The quantitative results which we have presented were reached only after a difficult analysis of a complex mass of data, which must now be described. It will be recalled that subjects were asked initially to judge which variable 'looked the same shape' as the standard. Subjects first require to be divided into two groups in terms of their initial reaction to these instructions: one group queried the instructions; the other group accepted them.

*Group querying instructions (n = 15)*

These subjects, before making a judgment, asked what the instructions meant. Each implied spontaneously that he could either make a 'real' judgment, or something contrasted with this. The phraseology might differ, but in every case it was clear that the alternative was 'real' or 'other-than-real'. These subjects were distributed as follows: 0°, nil; 25°, nil; 50°, four; 70°, five; 80°, six. Thus at smaller inclinations the instructions are not questioned; at greater inclinations some, but not all, question the instructions; and there is a suggestion that the greater the tilt, the more likely it is that the instructions will be questioned. These subjects were asked to make their preferred judgment, and were distributed as follows: 0°, nil; 25°, nil; 50°, 3 R/1 N; 70°, 4 R/1 N; 80°, 3 R/3 N. Thus a majority preferred the R judgment: 10 R/5 N. Each was then asked to make his second, non-preferred judgment. It should be particularly noted that every effort was made to avoid any suggestion that there might be more than one judgment, and any suggestion as to the nature of the judgments.

*Group accepting instructions (n = 85)*

The remaining eighty-five subjects chose one of the comparison triangles without questioning the instructions. They were next asked: 'Please describe what you have been doing'. It then became clear that they required to be divided into two groups. (1) One group ( $n = 23$ ) showed in their replies, spontaneously and promptly, that they were aware of the possibility of making both R and N judgments. They will be termed Group RN. In this group, some subjects had in fact made the R judgment and some the N judgment; and they were distributed as follows: 0°, nil; 25°, 1 R/1 N; 50°, 4 R/0 N; 70°, 6 R/2 N; 80°, 6 R/3 N. Thus a majority again preferred the R judgment: 17 R/6 N. Each was then asked to make his second, non-preferred judgment. It should again be emphasized that every effort was made to avoid any suggestion that there might be more than one judgment, and any suggestion as to the nature of the judgment. (2) The remainder ( $n = 62$ ), in their replies, showed awareness of only one possible judgment; further, all indicated, spontaneously and promptly, that they had in fact made R judgments. Since they were aware of the R judgment only, they will be termed Group RO.

*Groups RN and RO*

It will be clear that the fifteen subjects referred to above, who queried the instructions, may be grouped with the twenty-three subjects in Group RN, since all were spontaneously and promptly aware of the possibility of making both R and N

judgments. The total hundred subjects are therefore now divided as follows: Group RN ( $n = 38$ ), spontaneously aware of both judgments; Group RO ( $n = 62$ ), spontaneously aware of the R judgment only. Further, subjects in Group RN have made both R and N judgments, though some ( $n = 27$ ) made R judgments first, and some ( $n = 11$ ) N judgments first; while subjects in Group RO have made R judgments only. No subject is aware of the N possibility only; indeed, this judgment is described by contrast with the R judgment.

The distribution of subjects between the two groups is given in Table 2 ('Distribution between groups RO and RN'). At  $0^\circ$ , all subjects fall in group RO. As inclination increases, more subjects fall in group RN, and at  $70^\circ$  and  $80^\circ$  a majority fall in group RN. Thus at all inclinations all subjects are aware of the R possibility, but, as inclination increases, more subjects are likely to be aware of the N possibility also, though even at  $80^\circ$  one-quarter still fall in group RO.

Table 2 ('Choice of judgment') gives the numbers choosing each type of judgment at each inclination. A majority choose the R judgment at each inclination, but the number choosing the N judgment increases with inclination.

Table 2. Grouping of subjects

	Angle of inclination									
	0°		25°		50°		70°		80°	
Distribution between groups RO and RN										
Group RO	20		18		12		7		5	
Group RN	0		2		8		13		15	
Choice of judgment										
	R	N	R	N	R	N	R	N	R	N
Group RO	20	0	18	0	12	0	7	0	5	0
Group RN	0	0	1	1	7	1	10	3	9	6
Totals	20	0	19	1	19	1	17	3	14	6
Effect of instructions on Group RO										
Group RO	20		18		12		7		5	
Able to make N	—		1		3		4		3	
Ability to make the two judgments										
Total R	20		20		20		20		20	
Total N	0		3		11		17		18	
Choice of successive and simultaneous N judgments										
Nsc	0		3		10		15		16	
Nsm	0		0		1		3		2	
Sophistication										
Group RO	—		8/18		3/12		3/7		0/5	
Group RN	—		2/2		4/8		5/13		7/15	

### Further findings

(1) Having obtained these spontaneous judgments, an attempt was next made, by instructions and discussion, to make subjects in group RO aware of the N possibility, at inclinations  $25^\circ$  to  $80^\circ$  inclusive. This was successful in eleven of the total forty-two cases. Table 2 gives the distribution of these eleven ('Effect of instructions on Group RO'), and it will be seen that instructions are increasingly likely to be effective as inclination increases, since an increasing proportion of group RO can notice it.



(2) Table 2 ('Ability to make the two judgments') gives the total number able to make the R judgment (i.e., all subjects); and the total number able to make the N judgment (i.e. group RN plus the eleven subjects from group RO just mentioned). As inclination increases, subjects are increasingly likely to be able to make the N judgment; but even at  $80^\circ$  a few can still only make the R judgment.

(3) Table 2 ('Choice of successive and simultaneous N judgments') gives the numbers choosing the two forms of comparison for the N judgment. (At  $70^\circ$ , the numbers do not sum to the total able to make N judgments, because one subject used both methods of comparison.)

(4) No subject was acquainted with the relevant psychological literature; but some subjects might have knowledge of the principles of perspective from other sources. It might be suggested that such knowledge would predispose a subject to fall into group RN, and even that such knowledge might entirely account for the existence of group RN. Subjects were questioned on this point, and Table 2 ('Sophistication') gives the numbers in each group who possessed any marked knowledge of perspective. The data provide some support for the view that such subjects are more likely to fall into group RN. But it is clear that this does not provide a complete explanation of the occurrence of this group: some subjects in group RN do not possess this knowledge; some subjects who do possess the knowledge fall into group RO.

(5) The final data given in Table 1 include, for R judgments, the results of all subjects, and for N judgments, the results of all subjects able to make the N judgment, whether spontaneously or when instructed.

(6) Lastly we note that the distinction between groups RO and RN should not be drawn too rigidly. In group RO, some subjects never saw the possibility of making the N judgment; others achieved it after discussion. In group RN, some commented on the two possibilities only when asked to describe their judgments; others commented immediately. So there seems to be a continuum of readiness to notice the N possibility.

### *Summary*

The main finding is that one group of observers (RO) can initially only make R judgments, while a second group (RN), who increase with inclination, can make both R and N judgments. At  $0^\circ$  inclination, only the R judgment is possible. At the remaining inclinations, all subjects are still able to equate for 'real' shape. These judgments show comparatively little change, either in mean value or scatter, and are fairly accurate. A majority of subjects understand the phrase 'look the same shape' as requiring the R judgment at every inclination, but this interpretation becomes somewhat less 'salient' as inclination increases.

The N judgment seems to become much more 'salient' with inclination. Three pointers to this conclusion are: first, the number who spontaneously mention it increases; secondly, of those who mention it, an increasing proportion choose to make it; thirdly, an increasing proportion can notice it when instructed. None can make it at  $0^\circ$ , but almost all can make it at  $80^\circ$ . N judgments typically give a 'compromise' and are made with successive comparison; a few subjects elect to make a simultaneous

comparison and this gives a closer approximation to 'retinal' shape. Prior sophistication may partly account for the occurrence of N judgments, but seems unlikely to account for it entirely.

#### IV. REPORTS OF SUBJECTS

In asking subjects to describe their judgments, every effort was made to avoid leading questions. They used a variety of phrases, and individual differences are undoubtedly present. But there remains substantial agreement on certain main features. *Group RN* ( $n = 38$ ) will be considered first. (In this section, all phrases in inverted commas are quotations from subjects' reports.)

(1) The R judgment is concerned with 'what it is', 'the actual shape which is on that paper', 'the real shape'. Subjects were attempting to pick out a shape which could be exactly superimposed on the standard, and almost all were confident that the one they chose would lie within two or three steps of the standard. The N judgment is concerned with 'the shape as it looks, not as it is', 'as it appears to me now', 'the apparent shape'. Subjects contrast this with the R judgment; they describe it as something which they consider departs from the real shape.

(2) All subjects were well aware that the standard was tilted, and they regarded this as complicating the situation. There were several spontaneous comments, particularly at the extreme tilt: '*That angle... my goodness*'.

(3) Subjects typically spoke of the R judgment as 'taking the slope into account', 'considering the slope', or 'making allowance for the tilt'. This clearly suggests an activity on the part of the subject. But it is important to note that subjects' reports do not suggest that this is a very conscious or rational process. 'It is not exactly an intellectual allowance.' 'It is something you do instinctively as a matter of course in everyday life, and you are just applying the same sort of unconscious action to this situation.' The N judgment on the other hand is described as 'not taking account of the slope', 'ignoring the slope', or 'not making an allowance for the tilt'.

(4) Subjects state that it is necessary to allow for the tilt, when making the R judgment, because the effect of tilting the standard is 'to foreshorten it'; 'the whole thing will appear flatter than it actually is'; 'from this angle it looks smaller'. 'I should have thought it was obvious', said a subject at 50°, 'that the more it slopes away, the shorter it looks'. The N judgment is concerned with this foreshortened shape, which is described as a direct visual impression; it is 'as I actually see it', 'the immediate impression of direct shape', 'just looking'.

(5) Most subjects have no further analysis of the R judgment; but ten of the thirty-eight volunteered that 'allowing for the slope' implied imagining what the standard would look like if it were vertical. 'I try to think what it would look like if it were upright'; 'I have been turning it into a vertical plane'. (This seems analogous to 'potential looking' in the case of size; Joynson, 1958*b*.) In the case of the N judgment, on the other hand, it is often stated that it is based on the immediate present appearance: it is 'as that looks to me now'; 'just going by what it looks like now'.

(6) The two judgments are often described by the contrast between 'as it is' and 'as it looks'. But it is important to notice that many subjects nevertheless use the word 'look' in referring to the R judgment: it is 'when they look the same real shape',

or 'when they look the same allowing for the tilt'. We also saw that seventeen of these thirty-eight subjects initially took the phrase 'look the same shape' as requiring the R judgment without question.

(7) There is some evidence that the R judgment is the more natural and everyday. At every inclination a majority chose to make the R judgment first, either spontaneously or in response to the request to make their preferred judgment. Also, when questioned, thirteen subjects regarded the R judgment as more natural, and only seven the N judgment (this inquiry was incomplete). The data also hint that, as inclination increases, subjects are likely to become more evenly divided in this respect: this is suggested both by the changing proportions making the two judgments (Table 2), and by the subjects' reports.

(8) The Nsc and Nsm judgments are alike described as concerned with a direct visual impression. They seem to differ only in the method by which the comparison is made: namely, to adopt either a successive comparison, or a simultaneous comparison by juxtaposition.

To summarize, the *N judgment* is described as an attempt to equate on the basis of a direct, instantaneous, visual impression. This impression is held to depart from the real shape in consequence of the inclination at which the standard is presented, and the N judgment avoids allowing for this inclination. The *R judgment* is not a simple equation for direct visual impression. It allows for the slope in order to achieve an identical real shape, and is believed to be fairly accurate. Some subjects state that they are attempting to judge what the standard would look like if it were vertical. There is some evidence that the R judgment is the more natural.

For *Group RO* ( $n = 62$ ), we may begin at  $0^\circ$ . Here all twenty subjects fall in this group, for no ambiguity arises. 'Looks the same' and 'is the same' are indistinguishable. Further, subjects were aware that the standard was not tilted, and the question of allowing for the tilt did not arise.

When the standard was tilted, subjects in this group described the R judgment in the same way: 'looks the same' and 'is the same' remain identical. There was no doubt that subjects were aware that the standard was tilted, but there were no reports of 'allowing for the tilt'. In so far as subjects described how they made these judgments at all, it was superficial: 'I am judging angles to achieve an appearance of similarity'.

An attempt was next made, as we indicated earlier, to lead the forty-two subjects at inclinations  $25^\circ$  to  $80^\circ$  to see the distinction. They were told that some subjects found it necessary to allow for the slope because the triangle was foreshortened by its inclination. Eleven subjects then distinguished. They now described R and N judgments in the same way as Group RN. They said that they were not originally aware that they were 'allowing for the tilt', but thought they must have allowed unconsciously: 'I think I allowed without thinking about it'. The remaining thirty-one subjects could see no distinction, and also denied that they were allowing for the tilt. They either said that they thought inclination made little or no difference to the appearance; or that it simply made the judgment more difficult without specifying why. Finally, seven of these subjects made the rather striking statement that inclination, so far from foreshortening the standard, actually tended if anything to increase its apparent height. Thirty-one subjects, then, can see no important distinction



between 'looking the same' and 'being the same', even when the distinction is explained to them. The 'apparent' shape, and the 'real' shape, are identical. So for these subjects, the R judgment seems to approximate closely to that quality of direct, visual impression which for Group RN characterized the N judgment.

Some further findings emphasize this point. These thirty-one subjects were again pressed to distinguish the two judgments. Nineteen did not respond, but the remaining twelve behaved as follows. They concluded that the judgment which they had made was not in fact a judgment of real shape. Their argument was: 'I am told that a judgment of real shape involves allowing for tilt, and differs from a judgment of apparent shape. But my judgment did not allow for tilt, and was concerned with apparent shape. So I was mistaken in thinking I was judging real shape: I was simply judging apparent shape. I failed to realize that this is foreshortened by inclination, and I must make a correction to achieve the real shape.' They then made corrected estimates of real shape. At 25°, four subjects chose a larger triangle (mean value = 25.50), and one subject chose a smaller triangle (19.0). At 50°, five subjects chose a larger triangle (mean value = 24.10), and two subjects chose a smaller triangle (mean value = 18.0). Questioning showed that the three subjects who chose smaller triangles did so because they became muddled about the direction in which the correction was necessary.

Table 3. *R judgments for Groups RO and RN compared*

Measure	Group	Angle of inclination									
		0°	<i>n</i>	25°	<i>n</i>	50°	<i>n</i>	70°	<i>n</i>	80°	<i>n</i>
Mean	RO	22.40	20	21.61	18	21.83	12	21.71	7	18.60	5
	RN	—		21.50	2	22.38	8	22.00	13	21.60	15
Scatter ( $\sigma$ )	RO	1.39		1.16		1.28		1.83		1.74	
	RN	—		0.50		2.60		2.15		2.09	

It seems clear that all these corrected judgments are experimental artefacts: they are made only after long discussion with the experimenter; their quantitative value is atypical; and they have a deliberate and reasoned quality not found in other R judgments. They are of interest in showing the strength of the subject's conviction that his original judgment was a matter of direct, visual impression; and also as indicating one way in which marked 'over-constancy' can be obtained.

A last point concerns the *comparison of R judgments* for the two groups. The mean value and scatter are given separately in Table 3. At 25°, the number of subjects in Group RN is too small for valid comparison. At 50° to 80°, group RN is consistently greater in both mean values and scatter. The difference in mean values reaches significance at 80° ( $t = 3.05$ ;  $P < 0.01$ ). The difference in scatter reaches significance at 50° ( $F = 4.32$ ;  $P < 0.05$ ). Thus if the subject is aware of the N possibility, he seems likely to show rather greater 'constancy', than if he is not aware.

### Summary

The main point which emerges is that, for Group RN, the N judgment is an attempt to equate on the basis of a direct, instantaneous, visual impression; while the R judgment involves 'allowing for the slope'. For Group RO, however, the R judg-

ment approximates to a direct visual impression; though it differs from the N judgment as made by group RN both in its quantitative value, and in the fact that it is believed to be an equation for real shape.

The broad distinction between R and N judgments requires elaboration, however, to cover six main varieties of judgment.

(1) The R judgment at  $0^\circ$  tilt (Group RO): 'looks the same' and 'is the same' are indistinguishable; 'allowing for tilt' does not arise.

(2) The R judgment at  $25^\circ$  to  $80^\circ$  tilt (Group RO): described as (1); but some subjects may later distinguish and say that they unconsciously allowed for the tilt.

(3) The R judgment at  $25^\circ$  to  $80^\circ$  tilt (Group RN): 'allowing for the tilt'; contrasted with a direct visual impression.

(4) The Nsc judgment at  $25^\circ$  to  $80^\circ$  tilt (Group RN): a direct visual impression; contrasted with real shape; 'not allowing for the tilt'.

(5) The Nsm judgment at  $50^\circ$  to  $80^\circ$  tilt (Group RN): described as (4); but technique of juxtaposition employed.

(6) The corrected R judgment (Group RO): an artificial judgment, made when a subject concludes his original R judgment is incorrect, and usually giving overconstancy.

## V. DISCUSSION

The main finding is that one group (RO) can initially only make R judgments, while a second group (RN), who increase with inclination, can make both R and N judgments. We suggest that an interpretation is to be sought along the following lines.

In the everyday perception of shape, we are pre-occupied with 'real' shape. We are not usually concerned with the process by which our impressions of this are achieved. We take the process for granted, and concentrate on the result. If the question of the process is raised, we may well think that there is little to report, that the impression is simply 'given'. This is a first 'frame of mind', or 'attitude'. But in fact the impression involves a skill, and we may become aware of this and attempt to analyse it. We may, particularly at considerable tilts, note the contrast between the projective shape and what we judge to be the true shape. We now describe the projective shape as simply 'given', and the judgment of true shape as involving an activity. This is a second frame of mind. The first frame of mind might be termed '*unreflecting*', and the second '*investigating*'. It must be remembered that these terms refer to the subject's awareness of, or attitude to, his basis and method of judgment. They do not refer to the objects with which the judgment is concerned.

We suggest that in our data group RO approximate closely to the unreflecting frame of mind. They are pre-occupied with real shape. They take their skill for granted, and cannot analyse what it involves. If questioned, many think the impression is merely a direct presentation. Group RN represent the investigating state; their skill is more explicit. They are aware of the active nature of the real judgment, and contrast it with the immediately presented projective shape. Their greater awareness of what is involved is accompanied by slightly higher constancy in the real judgment. They predominate at greater tilts, presumably because the contrast between real and projective shape is greater there.

Whatever the validity of these suggestions, it is clear that results are conditioned

by the instructions, the experimental situation, and the subjects; and also that there is a complex interaction between these factors. If vague instructions are used, to say when things 'look the same', subjects may interpret them in different ways. In our experiment, this is in part a function of inclination. Variation in the angular separation of the objects might be expected to have a similar effect (cf. Joynson, 1958*a, b*). If subjects are instructed to judge 'real' shape, good constancy may be expected on the basis of the present results. If they are instructed to judge 'apparent' shape, more complex results may be anticipated. Some subjects may distinguish this from a judgment of real shape, and give N judgments; others may not distinguish, and give R judgments. The conditions of the experiment will influence the proportions falling in each group, and it will be essential to ask subjects what they have been doing. As we have found that subjects behave differently in identical situations, it is clear that factors within the subject are also significant. The experiment does little to elucidate these. Deep-seated differences in personality, intelligence, or past experience may be involved. More superficial factors, such as response to an experimental situation as such, or to a particular experimenter, might also be important.

Previous findings will be briefly considered from the point of view stressed, namely, that variations in conditions and procedure may be expected to influence results.

The early studies of Eissler (1933) and Klimpfinger (1933) concluded that results may be expected to differ according to whether the subject is 'object-directed' (synthetic attitude) or 'projective-directed' (analytic attitude). Interpretation of this requires caution, for these studies used few subjects (seven and five), and some of them were highly sophisticated (notably Brunswik). These attitudes do *not* seem to correspond to our unreflecting and investigating frames of mind. Rather we think that these sophisticated subjects were probably all in the investigating state; that the two classes of judgment correspond roughly to our R and N judgments as made by Group RN; and that the descriptions which they give of these, since they are incorporated in the instructions, are not to be regarded as experimental findings. Stavrianos (1945) instructed her five subjects to judge 'real' shape. The judgments obtained clearly correspond to our R judgments. In agreement with our study, good constancy resulted at the inclinations examined, namely from 15° to 55°. Leibowitz, Waskow, Loeffler & Glaser (1959) state that 'the subject is being asked to adopt an analytical attitude'. The resulting compromise judgments seem clearly to correspond to our N judgments. Sheehan (1938) instructed her twenty-five subjects to judge 'apparent' shape. Tilts of 15°, 45° and 75° were examined, though the order of presentation was not controlled. She finds constancy at first, with an increasing tendency to compromise as inclination increases. Unfortunately no attempt is made to determine what judgments subjects were making in response to these vague instructions. We can only suggest that Sheehan has combined R and N judgments indiscriminately. Her results are similar to what we should have obtained if we had combined all the first responses made, and then stopped the experiment. There is also a suggestion of bi-modality in the data. It seems advisable to abandon the term 'apparent shape'.

In the above experiments, the apparatus permitted the full range of possible judgments. But a number of experimenters (Thouless, 1931, 1932; Moore, 1938; Langdon, 1951, 1955; Lichte, 1952) employ apparatus in which the comparison series



does not contain the same real shape as the standard. This of course eliminates the possibility of R judgments, and it is only to be expected that, in all these experiments, compromise results are obtained. These experiments are therefore primarily concerned with what we have termed N judgments; and any interpretation of them requires to bear this limitation in mind.

It seems therefore that previous studies tend to give partial accounts of this subject. Their limitations must be remembered when they are interpreted. Unless we are mistaken, the present study is the first to examine the main possible judgments, spontaneously elicited from an adequate number of unsophisticated subjects, over the full range of inclinations. It is supported by the similar findings of a comparable experiment on size (Joynton, 1958*a*, *b*; Joynton & Kirk, 1960).

## VI. A SUBSIDIARY EXPERIMENT

The experiment was repeated on ten further subjects at 80°, using a modified method. Instead of the series of comparison triangles, subjects were presented with a sheet of paper on which was drawn a line the same length as the base of the standard triangle. They were asked to draw the remaining two sides to form a triangle which would 'look the same shape' as the standard. 10/10 subjects spontaneously distinguished between R and N judgments, as compared with 15/20 in the original experiment at 80°. 9/10 chose to make the N judgment first, as compared with 6/20 in the original experiment. Mean value for R judgments was 23.6 ( $\sigma = 5.04$ ); mean value for N judgments was 7.7 ( $\sigma = 4.27$ ). (All N judgments were Nsc not Nsm.) The R judgments are not significantly different from previous R judgments at 80° ( $t_{FB} = 1.5$ ;  $P > 0.05$ ), but scatter is significantly greater ( $F = 4.33$ ;  $P < 0.01$ ). The Nsc judgments are not significantly different from previous Nsc judgments ( $t = 1.08$ ;  $P > 0.05$ ), and scatter is not significantly different ( $F = 1.50$ ;  $P > 0.05$ ). The difference between mean values for R and Nsc judgments is however significantly greater in the present experiment than in the original experiment ( $t = 3.42$ ;  $P < 0.01$ ). This method thus probably (1) increases the likelihood of noticing the N possibility, (2) increases the preference for the N possibility, and (3) enhances the quantitative difference between the two judgments. It is suggested that a situation in which subjects are required to draw makes them more likely to think about perspective, with the above effects. Previous conclusions concerning the accuracy with which real shape can be judged at 80°, and concerning the spontaneity with which the N possibility is noticed, are confirmed.

## REFERENCES

- EISSLER, K. (1933). Gestaltkonstanz der Sehdinge bei Variation der Objecte und ihre Einwirkungsweise auf den Wahrnehmenden. *Arch. ges. Psychol.* **88**, 487-550.  
 GOTTHEIL, E. & BITTERMAN, M. E. (1951). The measurement of shape constancy. *Amer. J. Psychol.* **64**, 406-8.  
 JOYNSON, R. B. (1949). The problem of size and distance. *Quart. J. Exp. Psychol.* **1**, 119-35.  
 JOYNSON, R. B. (1958*a*). An experimental synthesis of the Associationist and Gestalt accounts of the perception of size. Part I. *Quart. J. Exp. Psychol.* **10**, 65-76.  
 JOYNSON, R. B. (1958*b*). An experimental synthesis of the Associationist and Gestalt accounts of the perception of size. Part II. *Quart. J. Exp. Psychol.* **10**, 142-54.

- JOYNSON, R. B. & KIRK, N. S. (1960). An experimental synthesis of the Associationist and Gestalt accounts of the perception of size. Part III. *Quart. J. Exp. Psychol.* **12**, 221-30.
- KLIMPFINGER, S. (1933). Über den Einfluss von intentionaler Einstellung und Übung auf die Gestaltkonstanz. *Arch. ges. Psychol.* **88**, 551-98.
- LAMBERCIER, M. (1946). La constance des grandeurs en comparaisons sériales. *Arch. de Psychol.* **31**, 79-282.
- LANGDON, J. (1951). The perception of a changing shape. *Quart. J. Exp. Psychol.* **3**, 157-65.
- LANGDON, J. (1955). The role of spatial stimuli in the perception of shape. Parts I, II. *Quart. J. Exp. Psychol.* **7**, 19-27; **7**, 28-36.
- LEIBOWITZ, H., WASKOW, I., LOEFFLER, N. & GLASER, F. (1959). Intelligence level as a variable in the perception of shape. *Quart. J. Exp. Psychol.* **11**, 108-12.
- LICHTE, W. H. (1952). Shape constancy: dependence upon angle of rotation; individual differences. *J. Exp. Psychol.* **43**, 49-57.
- MOORE, W. E. (1938). Experiments on the constancy of shape. *Brit. J. Psychol.* **29**, 104-16.
- SHEEHAN, M. R. (1938). A study of individual consistency in phenomenal constancy. *Arch. Psychol.* **31**, 1-95.
- STAVRIANOS, B. K. (1945). The relation of shape perception to explicit judgements of inclination. *Arch. Psychol.* **41**, 1-94.
- THOULESS, R. H. (1931). Phenomenal regression to the 'real' object. Parts I, II. *Brit. J. Psychol.* **21**, 339-59; **22**, 1-30.
- THOULESS, R. H. (1932). Individual differences in phenomenal regression. *Brit. J. Psychol.* **22**, 216-41.
- VERNON, M. D. (1952). *A Further Study of Visual Perception*. Cambridge University Press.

*(Manuscript received 6 February 1961)*





## THE EFFECTS OF MEPROBAMATE ON KINESTHETIC FIGURAL AFTER-EFFECTS

By C. G. COSTELLO

*Institute of Psychiatry (Maudsley Hospital), University of London*

The effects of 800 mg. meprobamate on kinesthetic figural after-effects were investigated. There was a tendency for subjects to overestimate the test block width during pre-stimulation trials. This tendency was found to be negatively correlated with extraversion as measured by the M.P.I. (1959). Meprobamate reduced the tendency of subjects to overestimate the test block width during pre-stimulation trials and also reduced the size of the figural after-effects produced by rubbing a stimulus block wider than the test block. Results for ascending and descending trials were analysed separately and the above effects were found only in the case of the descending trials. The effect of the interpolated stimulus block in producing figural after-effects is discussed in terms of the changing of the frame of reference set up between the test block and comparison block. Extraversion was not significantly correlated with any of the measures of figural after-effect.

### I. INTRODUCTION

Kinesthetic figural after-effects have been used to test predictions from theories of cortical functioning. These predictions have been related to individual differences (Eysenck, 1955; Nicholls, 1955; Wertheimer & Wertheimer, 1954) and to the effects of brain damage (Klein & Krech, 1952; Jaffe, 1954). The results are conflicting owing partly, it is felt, to lack of data as to what is actually happening in the test situation. This lack of basic information on the test situation probably accounts in part for the conflicting evidence with regard to the effects of drugs on kinesthetic figural after-effects.

Wertheimer, Levine & Wertheimer (1955) report that 'metabolic decreaseers'—100 mg. pethidine hydrochloride, 200 mg. quinalbarbitone, and 60 cc. of whisky—all lowered the perceptual measures used in their experiment. One of their perceptual measures was the kinesthetic figural after-effect. Though insufficient details are given in their paper as to the effects of the 'metabolic decreaseers' on the KFAE, it is apparent that they reduced the amount of the KFAE. The authors point out that this finding is consistent with their theory that the metabolically efficient individual will be more modifiable on perceptual measures such as the KFAE and will therefore show larger after-effects than the less metabolically efficient. Their finding is inconsistent with Eysenck's theory of cortical inhibition-excitation (1955) which would predict an increase in KFAE with the administration of 'metabolic decreaseers'. Briefly, Eysenck follows Kohler & Wallach (1944) in using the term 'figural after-effects' to denote the alterations which test objects may show when their figure currents pass through a satiated region—the satiation being produced, in the case of KFAE's, by rubbing the stimulus block. Eysenck equates satiation with reactive inhibition as used by Pavlov and Hull and the size of the KFAE is assumed to be proportional to the amount of satiation produced. Since the drug postulate of his theory states that depressant drugs will decrease excitatory potential and increase inhibitory potential, it is clear that drugs such as those used by Wertheimer would be predicted to increase the size of KFAE's.

Wertheimer *et al.* (1955) also found that 'metabolic increasers'—50 mg. ephedrine sulphate, 22 mg. strychnine sulphate, and 100 mg. nicotinic acid—did not produce a significant effect on the KFAE.

Poser, Knight & Lehmann (1958) reported that 265 mg. sodium amylobarbitone did not have a significant effect on the KFAE. They had predicted, on the basis of Eysenck's theory, that sodium amylobarbitone would increase the KFAE. They report that six out of their ten subjects showed decreases in the KFAE, which is in keeping with the findings of Wertheimer *et al.* (1955). They also report that 15 mg. of *d*-amphetamine sulphate resulted in a significant decrease in the KFAE, a finding in keeping with both theories mentioned above.

Eysenck & Easterbrook (1960) report that neither 5 mg. *d*-amphetamine sulphate, 90 mg. sodium amylobarbitone, nor 100 mg. meprobamate had a significant effect on the KFAE. They did, however, find that both sodium amylobarbitone and meprobamate increased the errors in the kinesthetic matching results.

It can be seen that the results of the investigations into the effects of drugs on KFAE's are conflicting. Neither Wertheimer's theory nor Eysenck's theory is supported by the evidence. In the investigation to be reported here, the effects of 800 mg. meprobamate on the KFAE was investigated. It was hoped that some light would be thrown on the effects of drugs on KFAE's by:

- (1) Analysing separately the results obtained for ascending and descending trials;
- (2) Considering an alternative explanation for the role of the stimulus block in producing KFAE's.

Meprobamate was chosen for the following reasons: It was desired to use a drug that acted as a mild depressant. Pfeiffer and his co-workers (1957) presented evidence, including studies of the effects of meprobamate on conditioned avoidance responses, strychnine and pentylene-tetrazol threshold and spinal reflex in animals and on the EEG of humans, which led them to conclude that meprobamate can be classified as a 'barbiturate-like drug'. Miletto, Colomb & Cardiare (1956), Tucker & Wilensky (1957), Henry & Obrist (1958), and Shagass, Azima & Sangowicz (1959) have all shown that meprobamate produced changes in the EEG similar to those produced by barbiturates. Further evidence in support of the use of meprobamate as a depressant is presented by Eysenck & Eysenck (1960). Five tests—nonsense syllable learning, reaction times, level of skin resistance to the passage of an electric current, flicker fusion and perimeter threshold differences—were administered under conditions of 'no drug', glutethimide, and meprobamate medication. In a canonical variate analysis of the results, only one significant latent root was obtained, suggesting that both meprobamate and glutethimide have depressant properties and are similar in their behavioural effects.

Meprobamate, although a depressant, does not appear to have such gross behavioural effects of sedation as barbiturates and is almost completely free of side-effects. Marquis and his co-workers (1957) found no deleterious behavioural effects of 800 mg. meprobamate in a study on motor skills, sensory processes, and judgment necessary for safe driving. Reitan (1957) found a significantly poorer performance only as a result of the administration of a large dose of meprobamate (1600 mg.) when compared with a placebo using a battery of tests selected from Halstead's battery (1947). Performance was not affected by 400 mg. meprobamate. Rokeach, Oram &

Marr (1959) found that neither analysis nor synthesis in thinking was affected for better or worse by 400 or 800 mg. meprobamate. Kornetsky (1958), using a multiple stimulus response apparatus, found that 1600 mg. meprobamate causes considerable motor impairment, whereas 800 mg. showed no effect. More recently, DiMascio & Rinkel (1960) and Klerman, DiMascio, Havens & Snell (1960) have reported that 800 mg. meprobamate increased speed of mental activity (serial addition), and that there was no impairment in psychomotor functions (tapping speed, visuomotor co-ordination and steadiness). Drowsiness, reported with quinalbarbitone and phenyltolozamine, was seldom reported with 800 mg. meprobamate.

Meprobamate then, though acting as a depressant, does not appear to have gross behavioural effects similar to those produced by barbiturates except when large doses are used. Since such effects may in some instances complicate the results of experiments, it is advantageous if they can be obviated. That the mild sedative effect of small doses of meprobamate may nevertheless be revealed by the use of some functions is suggested by the writer's earlier work (Costello, 1960), where it was shown that meprobamate produced significant changes in the after-image, the spiral after-effect, and apparent movement.

## II. SUBJECTS AND PROCEDURE

Twenty subjects were tested, twelve male and eight female. Their ages ranged from 17 to 47 years, with a mean of 29.95 years. All were paid volunteers and all were of average or above average intelligence.

Each subject was given two treatments—placebo and 800 mg. meprobamate on two different days, the order of treatments being counterbalanced. Both the meprobamate and placebo were in identically appearing tablet form and were taken orally with a draught of water. Testing was done 1½ hr. after administration of the treatment and approximately 2½ hr. after the last meal—a light lunch—had been taken.

The width of the test block was 1½ in., the stimulus block was 2½ in. and the comparison block was ½ in. wide at the near end, increasing in width to 4 in. at the far end. The apparatus is illustrated in Eysenck (1955).

The subject was blindfolded throughout the experiment. He was seated at a table on which were placed the test block and the comparison block. The test block and the stimulus blocks were always on the right of the subject and the comparison block always on the left. The comparison block remained in the same position with the narrower end always toward the subject.

The subject was instructed to grip the comparison block and test block using the thumb and big finger of each hand. He was instructed to keep the hand holding the test block still and to move the hand holding the comparison block to find the width on the comparison block that matched the width of the test block.

Four pre-inspection trials alternately ascending and descending were given to find the PSE and thus the constant error for each subject. The method of average error was used throughout the experiment.

The test block was then replaced by the 2½ in. stimulus block and the subject moved the thumb and big finger of his right hand back and forth along the sides of the stimulus block for a period of 15 sec. Immediately afterwards the subject made four judgments of the test block, alternately ascending and descending trials, and this was the first measure of FAE.

After a rest period of 1 min. the subject was given a stimulation period of 30 sec., followed by the second measurement of FAE. After a rest period of 5 min., four final estimates of the test block width were made.



III. RESULTS

The results for the ascending trials and for the descending trials were analysed separately. The results are shown in diagrammatic form in Figs. 1 and 2, and in Tables 1 and 2.

If we look at the results for the descending trials first of all we see that the drug has a significant effect on the pre-stimulation trials. The mean estimation on the placebo day is 10.07 (these scores are arbitrary units along the comparison scale, the point of objective equality to the test block being 9.0), and on the drug day it is 8.97; the difference between these means being significant at the 0.01 level. The

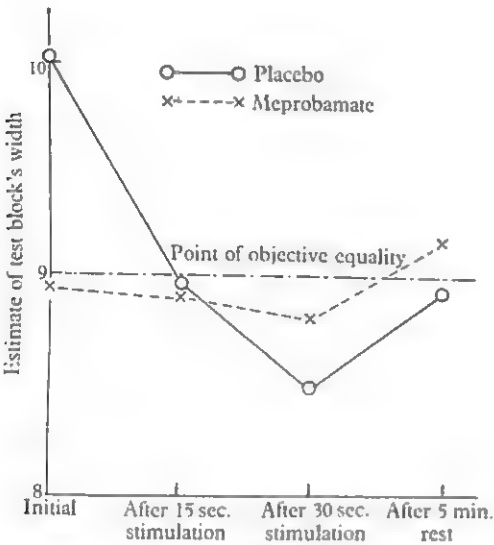


Fig. 1

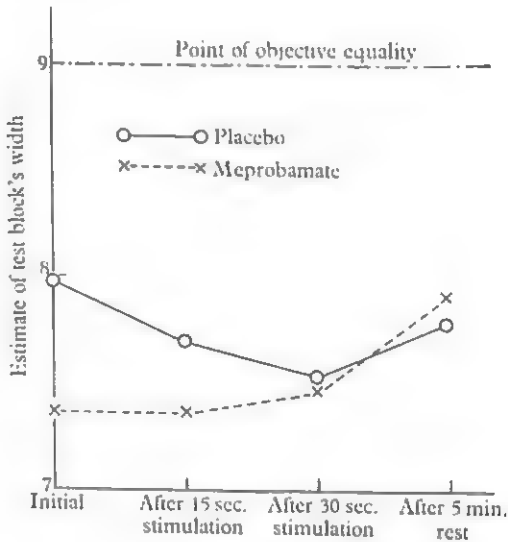


Fig. 2

Fig. 1. Estimates on descending trials (from wide end) of test block's width under placebo and drug treatments. Each point is the mean for twenty subjects. Point of objective equality to test block width represented by score of 9.

Fig. 2. Estimates on ascending trials (from narrow end) of test block's width under placebo and drug treatments. Each point is the mean for twenty subjects. Point of objective equality to test block width represented by score of 9.

Table 1. *Points of subjective equality on descending trials*

	Pre-stimulation trials		After 15 sec. stimulation		After 30 sec. stimulation		After 5 min. rest	
	M	S.D.	M	S.D.	M	S.D.	M	S.D.
Placebo day	10.07	1.56	8.96	1.69	8.47	1.26	8.92	1.54
Drug day	8.97	1.37	8.91	1.26	8.82	1.44	9.20	1.60

Table 2. *Points of subjective equality on ascending trials*

	Pre-stimulation trials		After 15 sec. stimulation		After 30 sec. stimulation		After 5 min. rest	
	M	S.D.	M	S.D.	M	S.D.	M	S.D.
Placebo day	7.98	1.41	7.67	1.15	7.51	1.25	7.76	1.58
Drug day	7.40	1.19	7.41	1.37	7.50	1.44	7.90	1.47

significance of the difference between means was tested by the *t*-test for matched groups and all the tests were two-tailed. The drug effect on the pre-stimulation estimates was expressed as a percentage of the placebo pre-stimulation readings since the difference between the placebo and drug readings was significantly correlated with the placebo readings ( $r = 0.64$ ,  $P < 0.01$ ). The percentage drop in the pre-stimulation readings produced by the meprobamate correlated insignificantly with the subject's extraversion score on the MPI (Eysenck, 1959) ( $r = -0.19$ ). There was, however, a significant negative correlation between extraversion and the pre-stimulation estimation on the placebo day ( $r = -0.51$ ,  $P < 0.01$ ).

The mean pre-stimulation estimate on the placebo day for the ascending trials was 7.98, on the drug day it was 7.40. The difference is in the same direction as for the descending trials but not statistically significant. Correlations of extraversion with the percentage drop produced by the drug ( $r = -0.02$ ) and with the pre-stimulation estimate on the placebo day ( $r = -0.19$ ) were statistically insignificant.

If we look again at the results for the descending trials, we find a significant drop in the estimation of the block width from the pre-stimulation trials to the trials after 15 sec. stimulation ( $P < 0.001$ ), a significant drop from the post-15 sec. stimulation trials to the post-30 sec. stimulation trials ( $P < 0.01$ ), and a significant rise from the post-30 sec. stimulation trials to the post-5 min. rest trials ( $P < 0.05$ ). None of these difference scores on the drug day were statistically significant. The drug reduces to insignificance the KFAE, which is in keeping with the results of Wertheimer *et al.* (1955) and the tendency found in the results of Poser *et al.* (1958).

Correlations were calculated between extraversion and the amounts of the KFAE on both the placebo day and the drug day. None of the correlations were significant.

In the case of the ascending trials, none of the measures of KFAE on either the placebo day or the drug day were significant. Correlations between extraversion and the amount of KFAE on both placebo day and drug day were also insignificant. As with the descending trials, there was a tendency for the drug to reduce the KFAE's.

#### IV. DISCUSSION

Let us consider the results we found on the pre-stimulation trials (descending). The test block in the right hand was overestimated on the placebo day, introverts overestimating the width to a greater extent than the extraverts. Meprobamate resulted in a slight underestimation of the test block width. Wertheimer (1954) reported that his right-handed subjects overestimated the size of a test block *held between the fingers of the left hand*. Wertheimer discusses his results in terms of an underestimation of the width of the comparison block held in the right dominant hand and notes that his results are in line with McPherson & Renfrew's finding (1953) that when physically equal-sized small metal circles are felt simultaneously by the two hands of a blind-folded subject the object held in the preferred hand is usually judged smaller. Since all the subjects used in the present investigation were right-handed, one would have expected an underestimation of the test block which was held in the right hand. As we saw, there was an overestimation of the test block. The results are obviously inconsistent with Wertheimer's hypothesis. It should be noted that Wertheimer's left-handed subjects also overestimated the test block *held in the left dominant hand*,



and underestimated the width of the comparison block held in the right non-dominant hand. Wertheimer comments that the number of left-handed subjects was too small to provide adequate data to test the effect but the tendency in the results is obviously inconsistent with his hypothesis that the important variable is the dominance or non-dominance of the hand. The results in Wertheimer's and the present study are more in line with the hypothesis that in pre-stimulation trials a test block width is overestimated as compared with that of a comparison block. Whether the dominance or non-dominance of the hand or the function of the block (test or comparison) is the important variable can only be adequately tested by giving groups of both right-handed and left-handed subjects trials with the test block in the left hand and trials with the test block in the right hand. Data on this problem are being collected at present.

Though there is no conclusive evidence that the difference in apparent width between a test block and a comparison block during pre-stimulation trials is due to the function of the block, the hypothesis has been elaborated further to account for the results obtained here. The further elaboration of the hypothesis is that the perceived width of a block of wood is proportional to the amount of pressure on the fingers holding the block. In the kinesthetic matching situation, it is assumed that more pressure is placed on the immobile fingers holding the test block than on the mobile fingers running along the comparison block.

It would follow from the above hypothesis that the degree of overestimation of a test block would be proportional to the difference in amount of pressure between the fingers on the test block and the fingers on the comparison block. A direct test of *this should not be too difficult*. How adequate is it in accounting for the data available? We may note that such an hypothesis might explain the significant difference found by Wertheimer (1954) between males and females in the overestimation of the test block. The disappearance of this overestimation after administration of meproamate, a C.N.S. depressant and muscle relaxant, is also what might be expected since the drug could be expected to reduce the pressure on the test block. The lesser tendency of the extraverts to overestimate the block might also result from the build up of satiation, thus reducing the pressure on the test block. It should be noticed that the explanation offered for the overestimation of the test block on the descending trials does not help us to understand the underestimation of the test block on the ascending trials. It may be that starting from the wide or narrow end has an effect on the pressure with which the test block is held or that the overestimation or underestimation is a result not only of pressure differences, as we have suggested, but also of the general frame of reference set up by starting at one end or the other of the comparison block.

Let us look now at the KFAE produced by stimulation. The general finding is that meproamate reduces the KFAE. Eysenck's theory, based on the assumption that the degree of KFAE is proportional to the degree of satiation produced by the stimulation, would predict an increase in KFAE produced by the drug. Wertheimer's theory based on the assumption that the degree of KFAE is proportional to the perceptual modifiability of the subject would predict a decrease in the KFAE. Before we immediately decide to accept Wertheimer's theory, however, let us look at further data as to what is happening on the kinesthetic figural after-effect test.



Jaffe (1956) has produced results showing the influence of visual stimulation on KFAE's. Twenty subjects were tested with apparatus similar to that used here, but the interpolated stimulus block was equal to the test block. Control subjects showed no KFAE but experimental subjects who looked at parallel lines equal to the test block during estimation of the test block width and wider or narrower than the test block during stimulation showed significant distortions after the wider stimuli. In a second experiment, subjects were given visual stimulation (wider or narrower than the test block) during the stimulus period and no visual stimulus during the testing period. In this case no KFAE appeared. Jaffe concludes that the results of both experiments 'can be summarized by saying that concomitant visual stimulation can induce a significant kinesthetic figural after-effect but that this after-effect appears only when contrasting size relationships exist between the visual stimuli'. He also writes that the difference between the two experiments 'suggests that the contrast between the visual objects resulted in a change in the general frame of reference and this in turn influenced the kinesthetic judgement'.

Jaffe's experiments suggest that in the usual KFAE test situation the interpolated wider stimulus block results in a decrease in the estimation of the test block width owing to a change in the general frame of reference.

Using this kind of explanation, we may argue that in the present experiment the test block was initially overestimated owing to differences in pressure on the blocks. The general frame of reference relating the comparison to the test block was then changed by the interpolation of the wider stimulus block. It was changed in such a way that the width of the test block was no longer overestimated. If the amount of the effect on the frame of reference of the interpolated stimulus block is related to the degree of retention of the effects of the interpolated stimulus block, then one would expect lower KFAE's after administration of a depressant drug. This suggests that the original assumptions regarding the test rather than the theory must be rejected. (If we can assume that the retention of the effects of the interpolated stimulus blocks is related to the continuation of excitation from the stimulus block, then the results can be predicted from the drug postulate of the theory of cortical excitation-inhibition.)

On the basis of the above hypothesis, one would expect that extraverts would show less KFAE. The correlations between extraversion and the amount of KFAE in this experiment were insignificant. Eysenck (1955) reported positive correlations between extraversion and the amount of KFAE but Broadbent has since shown that this may be due to the fact that the pointed edge of the wedge was always towards the subject in Eysenck's study and that the correlation may be between extraversion and a centralizing tendency. McEwen & Rodger (1960) failed to find any correlation between extraversion and the amount of KFAE in a situation designed to rule out the possible influence of centralizing tendencies. Nicholls (1955) did not obtain significant correlations between extraversion and the amount of KFAE. There was a tendency for a group of hysterics to experience larger KFAE's than a group of dysthymics but the differences were not significant.

It would also be predicted that brain-damaged subjects would have larger figural after-effects than normal subjects since Eysenck postulates that inhibitory potential is increased by brain damage. Klein & Krech (1952) reported that brain-damaged

subjects had larger KFAE's but their differences were statistically insignificant and Jaffe (1954) found no differences between the KFAE's of brain-damaged and normal subjects.

At present then the above discussion relating the theory of excitation-inhibition to KFAE must be restricted to the drug data. The adequacy of the theory to account for the effects of extraversion and brain damage on the KFAE's can only be decided when less equivocal data are at hand.

A word or two more should be said about Broadbent's hypothesis (1961). It may be argued that the greater overestimation of the test block in the present study on the placebo day pre-stimulation trials as compared with the drug day results may be due to the drug subject's bias towards the body. It may be argued also that the negative correlation between extraversion and the pre-stimulation estimates on the placebo day are due to the extravert's bias towards the body. But this is unlikely since the difference occurred on the first four trials whereas Broadbent's effect appears only after a larger number of trials and, according to Broadbent, occurs as a result of a greater familiarity with the comparison block.

Broadbent's further discussion of his hypothesis suggests that what he at first calls a 'bias towards the body' tendency is actually a greater centralizing tendency and a greater tendency to avoid extreme judgments. If this is so, then extraverts should have shown a greater tendency in the pre-stimulation trials to choose a point between the wide end of the comparison block and the point of objective equality. But as we have seen it was the introverts who showed this tendency.

Broadbent also showed that pseudo-KFAE's occurred without interpolated stimulation, which he felt were due to a centralizing tendency. In the case of subjects who had the narrow end of the wedge pointed towards them and were only given ascending trials, the PSE was shifted towards the narrow end of the comparison block. This apparently was due to the fact that they were only familiar with that part of the comparison block between the narrow end and the point of objective equality. In the case of subjects given descending trials (starting from the wide end), the PSE moved towards the wide end, resulting in an overestimation of the test block. This was due apparently to the broadening of the subject's experience of the block. Broadbent concludes that 'the tendency to avoid extreme judgments causes a trend in the PSE which will appear even in the absence of any distorting experience'.

It may be more correct to say that this tendency to avoid extreme judgments which causes a trend in the PSE may be overcome as a result of a distorting experience. In Broadbent's Groups II and III where an interpolated stimulus block was used, there was an underestimation of the test block even though ascending and descending trials were used. It would seem then that though a centralizing tendency may be operative (and in his Groups II and III this is correlated with extraversion) the interpolated stimulus block also plays a part in reducing the perceived width of the test block. The fact that the subjects in the present experiment underestimated the test block after interpolated stimulation and that this occurred on descending trials also indicates that interpolated stimulation may overcome any centralizing tendencies.

Broadbent discusses his correlation between extraversion and the centralizing tendency in terms of the reactive inhibition of judgments on early trials. He did not

find a correlation when no interpolated experience was used. He suggests that the interpolated experience causes forgetting of the first series of judgments but admits that his is an *ad hoc* assumption. It is felt that Broadbent relegates the interpolated experience to a too minor role, particularly in view of the results of Jaffe (1956) who found differences in the effects of visual stimuli of different sizes and indeed in view of the general finding of the differences between the effects of stimulus blocks wider or narrower than the test block.

The author agrees with Broadbent in so far as he considers changes in the PSE as being due to changes in the frame of reference. But it is felt that the width and amount of time spent on the interpolated block has a direct effect on the frame of reference rather than merely interfering with the frame of reference established by early trials. The author agrees with Broadbent's suggestion that the frame of reference hypothesis is in keeping with the permanent KFAE's found by Wertheimer & Leventhal (1958). It is certainly more in keeping with this finding than the original hypothesis of Eysenck (1955). It may be noted that Kohler & Dinnerstein (1947) also found that figural after-effects tended to become permanently established.

#### REFERENCES

- BROADBENT, D. E. (1961). Psychophysical methods and individual differences in the kinesthetic figural after-effect. *Brit. J. Psychol.* **52**, 97-104.
- COSTELLO, C. G. (1960). The effects of meprobamate on perception. *J. Ment. Sci.* **106**, 322-36.
- DiMASCIO, A. & RINKEL, M. (1960). Prediction of clinical effectiveness of psychopharmacologic agents from 'Drug action profiles' in normal human subjects. In: Wortis, H. (Ed.), *Biological Psychiatry*, vol. 2. New York: Grune and Stratton.
- EYSENCK, H. J. (1955). Cortical inhibition, figural after-effect and theory of personality. *J. Abnorm. Soc. Psychol.* **51**, 94-106.
- EYSENCK, H. J. (1959). *The Maudsley Personality Inventory*. London: University of London Press.
- EYSENCK, H. J. & EASTERBROOK, J. A. (1960). Drugs and personality. X. The effects of stimulant and depressant drugs upon kinesthetic figural after-effects. *J. Ment. Sci.* **106**, 852-7.
- EYSENCK, H. J. & EYSENCK, S. B. G. (1960). The classification of drugs according to their behavioural effects: a new method. In: Eysenck, H. J. (Ed.), *Experiments in Personality*. London: Routledge and Kegan Paul.
- HALSTEAD, W. C. (1947). *Brain and Intelligence: a Quantitative Study of the Frontal Lobes*. Chicago, Ill.: University of Chicago Press.
- HENRY, C. E. & OBRIST, W. (1958). The effect of meprobamate on the EEG. *J. Nerv. Ment. Dis.* **126**, 268-71.
- JAFFE, R. (1954). Kinesthetic after-effects following cerebral lesions. *Amer. J. Psychol.* **67**, 668-76.
- JAFFE, R. (1956). The influence of visual stimulation on kinesthetic figural after-effects. *Amer. J. Psychol.* **69**, 70-5.
- KLEIN, G. S. & KRECH, D. (1952). Cortical conductivity in the brain injured. *J. Personality*, **21**, 118-48.
- KLERMAN, G. L., DiMASCIO, A., HAVENS, L. L. & SNELL, J. E. (1960). Sedation and tranquilization: A comparison of the effects of a number of psychopharmacologic agents upon normal human subjects. *A.M.A. Arch. Gen. Psychiat.* **3**, 28-37.
- KOHLER, W. & DINNERSTEIN, D. (1947). Figural after-effects in kinesthesia. In: Michotte, A. (Ed.), *Miscellanea Psychologica*, pp. 196-220. Louvain: Publications Universitaires de Louvain.



- KOHLER, W. & WALLACH, H. (1944). Figural after-effects: an investigation of visual processes. *Proc. Amer. Phil. Soc.* **88**, 269-357.
- KORNETSKY, C. (1958). Effects of meprobamate, phenobarbital and dextroamphetamine on reaction time and learning in man. *J. Pharmacol.* **123**, 216-19.
- MARQUIS, D. G., KELLY, E. L., MILLER, J. G., GERARD, R. W. & RAPOPORT, A. (1957). Experimental studies of behavioural effects of meprobamate on normal subjects. *Ann. N.Y. Acad. Sci.* **67**, 701-11.
- MC EWEN, P. & RODGER, R. S. (1960). Some individual differences in figural after-effects. *Brit. J. Psychol.* **51**, 1-8.
- MC PHERSON, A. & RENFREW, S. (1953). Assymetry of perception of size between the right and left hands in normal subjects. *Quart. J. Exp. Psychol.* **5**, 66-74.
- MILETTO, G., COLOMB, H. & CARDIARE, G. (1956). Action du meprobamate sur L'E.E.G. humain. *Electroencephalog. Clin. Neurophysiol.* **8**, 715.
- NICHOLLS, E. G. (1955). The relation between certain personality variables and the figural after-effect. Unpublished Ph.D. thesis, University of London.
- PFEIFFER, C. C., RIOPELLE, A. J., SMITH, R. P., JENNY, E. H. & WILLIAMS, H. L. (1957). Comparative study of the effect of meprobamate on the conditioned response on strychnine and pentylenetetrazol thresholds on the normal electroencephalogram and on polysynaptic reflexes. *Ann. N.Y. Acad. Sci.* **67**, 734-45.
- POSER, E. G., KNIGHT, D. A. & LEHMANN, H. E. (1958). *The influence of excitant and depressant drugs on kinesthetic figural after-effect*. Read at Annual Convention of the A.P.A., 1958.
- REITAN, R. M. (1957). The comparative effect of placebo and meprobamate on psychologic test performance. *Antibiotic Med. Clin. Therapy*, **4**, 158-65.
- ROKEACH, M., ORAM, A. & MARR, J. N. (1959). The effects of meprobamate on analysis and synthesis in thinking. *J. Psychol.* **48**, 359-66.
- SHAGASS, C., AZIMA, H. & SANGOWICZ, J. (1959). Effect of meprobamate in sustained high dosage on the electroencephalogram sedation threshold. *Electroencephalog. Clin. Neurophysiol.* **11**, 275-83.
- TUCKER, K. & WILENSKY, H. (1957). A clinical evaluation of meprobamate therapy in a chronic schizophrenic population. *Amer. J. Psychiat.* **113**, 698-703.
- WERTHEIMER, M. (1954). Constant errors in the measurement of figural after-effects. *Amer. J. Psychol.* **67**, 543-6.
- WERTHEIMER, M. & LEVENTHAL, C. M. (1958). Permanent satiation phenomena with kinesthetic figural after-effects. *J. Exp. Psychol.* **55**, 255-7.
- WERTHEIMER, M., LEVINE, H. & WERTHEIMER, N. (1955). The effect of experimentally induced changes in metabolism on perceptual measures of metabolic efficiency. *Percept. Mot. Skills*, **5**, 173-6.
- WERTHEIMER, M. & WERTHEIMER, N. (1954). A metabolic interpretation of individual differences in figural after-effects. *Psychol. Rev.* **61**, 279-80.

(Manuscript received 1 June 1960)

## UNCERTAINTY AND EPISTEMIC CURIOSITY

By D. E. BERLYNE\*

*Boston University*

Two experiments were concerned with the effects of uncertainty on epistemic curiosity. Quotations, each coupled with the names of possible authors and with a distribution of fictitious experts' guesses regarding the true author, formed the experimental material. The role of uncertainty was predicted from a theory according to which epistemic curiosity will increase with degree of conceptual conflict.

The two experiments showed reported curiosity to increase with two determinants of uncertainty and degree of conflict, namely (1) number of alternative responses and (2) evenness of distribution of response-strength, respectively. They also revealed ways in which response uncertainties, dependent on distributions of subjects' guesses, differ from the corresponding stimulus uncertainties.

A theory of *epistemic curiosity*, conceived as a motivational state (a state of high drive or arousal) that actuates quests for knowledge and is relieved by acquisition of knowledge, has been outlined in a number of recent publications (Berlyne, 1954*a*, 1957, 1960*a, b*). According to this theory, the strength of epistemic curiosity will increase with the degree of *conceptual conflict* or conflict between symbolic response-tendencies—beliefs, attitudes or thoughts. Degree of conflict is, in its turn, assumed to increase with (1) the number of competing response-tendencies, (2) their total absolute strength, (3) their nearness to equality of strength, and (4) their degree of mutual incompatibility.

Some data confirming predictions from this theory were obtained in an earlier experiment (Berlyne, 1954*b*). In this experiment, one group of subjects received (1) a fore-questionnaire of 48 factual multiple-choice questions about invertebrate animals, followed by (2) a series of statements including answers to these questions, and finally (3) an after-questionnaire putting the same questions in an open-ended form. Immediately after attempting the fore-questionnaire, subjects were required to mark the three questions out of each consecutive set of 12 whose true answers they would most like to know.

The following findings emerged, among others:

(1) Questions that subjects marked in the fore-questionnaire as those to which they would most like to know the answers were more likely than others to be answered correctly in the after-questionnaire. Consequently, both of these operators (which we may call the Marking Test and the Retention Test respectively) were regarded as measures of epistemic curiosity.

(2) According to both the Marking Test and the Retention Test, questions about more familiar animals induced more curiosity than those about less familiar animals. This result was predicted on the assumption that symbolic responses associated with more familiar concepts will be both more numerous and greater in absolute strength.

(3) According to the Retention Test, more curiosity was induced by questions that subjects found surprising and by questions whose predicates appeared to a group

\* Now at the University of Toronto.

of judges to be least applicable to the animals concerned. The questions singled out by both of these criteria may be assumed to have juxtaposed concepts with a high degree of mutual incompatibility.

So, in brief, the hypothesized effects on epistemic curiosity of three of the four determinants of degree of conflict received some measure of confirmation. But the confirmation was not so conclusive or direct on all counts as could be desired. Furthermore, the remaining determinant, namely the nearness to equality of strength of competing responses, was not investigated at all.

The present experiments were therefore intended to carry the line of research further. The effects of degree of conceptual conflict were sounded through the information-theoretic measure, *uncertainty* (otherwise known as *entropy*). The relations between uncertainty and conflict seem to be close (Berlyne, 1957; 1960*a*, ch. 2). The situations in which uncertainty regarding stimuli or responses has proved to be a measure of psychological importance are all situations that can be supposed to entail appreciable conflict. Uncertainty also satisfies some, but not all, of the requisites for a measure of degree of conflict as a function of response-strengths. In particular, and this is of importance for our present inquiry, uncertainty goes up (1) as the number of alternatives increases and (2) as the alternatives approach equiprobability. Equiprobability can be regarded as a reflection of equal response-strength, so that uncertainty reflects two of our determinants of degree of conflict, including the one that was not investigated in the previous experiment.\*

## EXPERIMENT I

### *Subjects*

Data were collected from four eleventh-grade (i.e. about 16-year-old) high-school classes. Two classes, whose total membership was 40, formed the 'Guess' group. The other two classes, with a total membership of 44, formed the 'No-Guess' group.

### *Material*

A series of thirty quotations, one or two sentences long, from men prominent in American or English literature or history of the last two centuries was drawn up. The quotations were all highly unlikely to be familiar to the subjects.

Three alleged authors were selected to be associated with each quotation. In order to control for previous knowledge that might have facilitated the recognition of the true author, the alleged authors did not include the true author, but they were, as far as possible, men who could plausibly have been responsible for the quotation in question. All of them figured prominently either in current newspapers or in the subjects' text-books for courses on history and literature.

\* Whenever we have a situation where an event may fall into any of  $n$  alternative classes, and we can assign a probability,  $p$ , to each of these classes, we can calculate the value of a variable,  $H$ , according to the formula

$$H = - \sum_{i=1}^n p_i \log_2 p_i.$$

$H$  will then represent the 'uncertainty', 'entropy' or 'amount of information to be extracted' of the situation in question. The unit of  $H$  is generally called the 'bit', one bit being the uncertainty that obtains when an event may fall into either of two equally likely classes.



*Manipulation of uncertainty*

Uncertainty ( $H$ ) was manipulated as follows. Subjects were told at the start of the experiment that each of the quotations had been shown to a group of 100 high-school teachers of history and literature, together with the names of two or three possible authors. The teachers had been asked, it was stated, to indicate which they thought was the true author.

Each quotation was accordingly followed in the experimental booklet by the names of two or three alleged authors, each coupled with a number. Subjects were given to understand that the numbers represented how many teachers, out of 100, had selected each of the possible authors as the actual author.

For every subject, there were, in a random order, 10 high-uncertainty quotations, 10 medium-uncertainty quotations, and 10 low-uncertainty quotations, depending on the distributions of the alleged teachers' guesses:

(1) The *high-uncertainty* quotations had three alleged authors, and distributions clustering round (34, 33, 33), for which  $H = 1.58$  bits. The actual distributions varied slightly for different quotations, as the use of the same distribution 10 times would have destroyed credibility.

(2) The *medium-uncertainty* quotations comprised two subgroups. There were five with two authors and distributions clustering round (50, 50) and five with three authors and distributions clustering round (77, 13, 10). For both of these,  $H = 1.00$  bit.

(3) The *low-uncertainty* quotations had two authors and distributions clustering round (90, 10), for which  $H = 0.47$  bit.

Since  $H$  depends on two of the hypothesized determinants of degree of conflict, namely nearness to equality of response-strength and number of alternative responses, it was deemed advisable to split the medium-uncertainty quotations into two subclasses of five quotations, so that the effects of these two determinants could be assayed separately. The 30 quotations thus comprised 15 with three alternatives and 15 with two alternatives. They also comprised 15 with even distributions and 15 with uneven distributions.

There were three versions of the experimental booklet, each given to approximately equal numbers of subjects. What were high-uncertainty quotations in one version formed the medium-uncertainty set in another version and the low-uncertainty set in the remaining version, so that each quotation appeared in each of the three uncertainty sets about equally often.

*Procedure*

Subjects were first told that they were participating in a research project on ways of improving teaching methods. The present experiment, whose connexion with the aim of the project might not be evident, was concerned with ability to recognize authors of quotations. It was pointed out that experts could often tell who wrote or said something from what they knew about the person or from the nature of the quotation. But at times even an expert would be surprised to learn who the actual author was. The mendacious information about the guesses of the fictitious teachers was then furnished.

An experimenter then read out each of the 30 quotations in turn, while the subjects

followed the reading from the mimeographed version in the booklet. There was a pause of about 12 sec. between the readings of two successive quotations, during which the subjects were to look at the names of the possible authors and the number of teachers selecting each as the true author. Subjects in the Guess group were also required to underline whichever they thought was the name of the true author, while all that the No-Guess group had to do at this stage was to look at the written question while it was being read aloud.

When the reading was concluded, subjects were told to go back over the 30 quotations and mark the 12 quotations whose true authors they would most like to know. Finally, they were to go back over the quotations once more and rank-order the 12 that they had marked.

### Results

*Curiosity scores.* Scores positively related to amount of curiosity were obtained by allotting 0 to unmarked quotations and subtracting the ranks assigned to marked quotations from 13. The mean curiosity scores for the various groups of quotations were then calculated.

No significant differences between Guess and No-Guess groups appeared, and so the data for both groups will be discussed together.

As the third row of Table 1 shows, the mean curiosity score was greatest for high-uncertainty quotations, next greatest for medium-uncertainty quotations, and smallest for low-uncertainty quotations. According to the results of a Friedman non-parametric analysis of variance,  $\chi^2_r = 18.66$  for 2 D.F., and  $P < 0.001$ .

Table 1. *Experiment I*

	Low-uncertainty quotations	Medium-uncertainty quotations		High-uncertainty quotations
Stimulus distributions (%)	(90, 10)	(50, 50)	(77, 13, 10)	(34, 33, 33)
Stimulus uncertainty (bits)	0.47	1.00	1.00	1.58
Mean curiosity scores	2.17	2.67		3.02
		2.29	3.04	
Response distributions (%)	(72, 28)	(54, 46)	(58, 18, 24)	(35, 35, 30)
Response uncertainty (bits)	0.86	1.00	1.39	1.58

It would appear from this finding that curiosity increased with uncertainty, but the confidence that can be placed in this conclusion is severely attenuated by a comparison of the mean curiosity scores for the two medium-uncertainty subsets. This comparison, as can be seen from Table 1, reveals the mean curiosity score for the (77, 13, 10) quotations to be appreciably greater than that for (50, 50) quotations, and, in fact, it even exceeded the mean curiosity score for (34, 33, 33) quotations slightly. A sign test applied to the difference between the two subclasses of medium-uncertainty questions showed the (77, 13, 10) to have higher scores for 54 out of 84 subjects, with three ties, so that  $P < 0.01$ .

This makes it look as if one of the determinants of uncertainty, namely number of alternatives, was affecting curiosity, while the other determinant, evenness of distribution, was not. So it was decided to examine the effects of these two factors separately.

Orthogonal comparisons were contrived by giving double weight to the scores for quotations in the two medium-uncertainty subsets. The adjusted mean for three-alternative quotations then came to 2.84 and the adjusted mean for two-alternative quotations to 2.23. The difference was in this direction for 61 out of 84 subjects, with one tie, so that  $P < 0.01$  according to the sign test. The adjusted means for even-distribution and uneven-distribution quotations were 2.66 and 2.42 respectively, and although this difference was in the predicted direction for the experiment as a whole, it followed this direction in only 44 out of 84 subjects, with one tie, which did not approach significance.

So the results uphold our prediction that curiosity will be greater for three-alternative than for two-alternative quotations, but fail to confirm our prediction that curiosity will increase with evenness of distribution.

*Response uncertainties.* In order to throw further light on these results, it was decided to analyse the distributions of guesses in the Guess group with a view to ascertaining the *response uncertainties* for the different sets and subsets of quotations. Response uncertainty is, of course, what best provides a measure of degree of conflict. Differences in stimulus uncertainty were introduced solely as a means of generating differences in response uncertainty and thus differing degrees of conflict.

The fourth row of Table 1 displays the percentages of quotations for which the alleged authors mentioned first, second and third respectively were guessed to be correct by the Guess group, with the corresponding response uncertainties in the fifth row. The authors' names were always printed in decreasing order of alleged teachers' guesses, as this is how similar data will usually have been presented in the subjects' experience. A comparison of response uncertainties with stimulus uncertainties shows the following:

(1) Response uncertainties were affected by stimulus uncertainties. This was confirmed by statistical tests: the mean proportion of selections of the first author was greater for (90, 10) than for (50, 50) quotations ( $t = 3.84$ , 39 D.F.,  $P < 0.01$ ), and greater for (77, 13, 10) than for (34, 33, 33) quotations ( $t = 5.41$ , 39 D.F.,  $P < 0.01$ ).

(2) There was a tendency, when stimulus distributions were even, to favour the first author, but this did not prevent response uncertainties from coming very close to stimulus uncertainties in these cases.

(3) With uneven-distribution quotations, response distributions were markedly less uneven than stimulus distributions, so that, in these cases, response uncertainties were considerably higher than the corresponding stimulus uncertainties.

The upshot was that, whereas the response uncertainties differed markedly between three-alternative and two-alternative quotations, the difference in response uncertainty between even- and uneven-distribution quotations became very small. Since this latter fact may have been responsible for the lack of evidence that evenness of distribution affects curiosity, it was thought worth while to perform Expt. II with the same procedure but with greater differences in evenness of stimulus distribution than had been used in Expt. I.



EXPERIMENT II

Subjects

Data were collected from two discussion sections of the Introductory Psychology courses at the University of Maryland and from four eleventh-grade high-school classes. One discussion section (of 28 members) and two high-school classes (with a total of 35 members) comprised the Guess group, while the other discussion section (of 26 members) and the other two high-school classes (totalling 46 members) comprised the No-Guess group.

Material

The material was the same as in Expt. I, with the following exceptions. Only 28 quotations were used, divided into four sets of seven each:

- (1) The 3-even quotations had three authors and stimulus distributions clustering round (34, 33, 33), so that  $H = 1.58$  bit.
- (2) The 3-uneven quotations had three authors and stimulus distributions clustering round (97, 2, 1), so that  $H = 0.22$  bit.
- (3) The 2-even quotations had two authors and stimulus distributions clustering round (50, 50), so that  $H = 1.00$  bit.
- (4) The 2-uneven quotations had two authors and stimulus distributions clustering round (98, 2), so that  $H = 0.14$  bit.

The experimental booklet was prepared in two versions, each of which was given to half of the subjects in each group. The quotations that had three alternative authors in one version had two alternative authors in the other and vice versa, while those that had even distributions in one version had uneven distributions in the other and vice versa.

Procedure

The procedure was exactly as in Expt. I, except that subjects were simply told that the experiments were concerned with ability to recognize authors of quotations.

Table 2. *Experiment II*

	2-uneven quotations	2-even quotations	3-uneven quotations	3-even quotations
Stimulus distributions (%)	(98, 2)	(50, 50)	(97, 2, 1)	(34, 33, 33)
Stimulus uncertainty (bits)	0.14	1.00	0.22	1.58
Mean curiosity scores	2.43	3.26	2.69	2.96
Response distributions (%)	(75, 25)	(51, 49)	(66, 19, 15)	(33, 38, 19)
Response uncertainty (bits)	0.76	1.00	1.13	1.58

Results

Since the data for the university students and for the high-school students were strikingly similar, they will be discussed together. Their pooled results are shown in Table 2.

*Evenness of distribution.* For all 135 students, the mean curiosity score for even-distribution quotations was 2.98 and the mean curiosity score for uneven-distribution quotations was 2.56. According to a Wilcoxon signed-rank test for matched pairs,  $z = 2.54$ , and  $P$  (one tail)  $< 0.01$ .

The corresponding means for the 63 subjects in the Guess group were 2.87 and 2.67  $z = 1.76$ ,  $P < 0.05$ . For the 72 subjects in the No-Guess group, the means were 3.10 and 2.47:  $z = 2.71$ ,  $P < 0.01$ .

*Number of alternatives.* The overall means for three-alternative and two-alternative quotations were 2.82 and 2.76. The corresponding means for the Guess group were 2.86 and 2.71, and those for the No-Guess group were 2.77 and 2.80. None of these differences, of course, approached significance.

*Response uncertainties.* The response distributions for the Guess group and the corresponding response uncertainties were computed as in Expt. 1. They appear in the fourth and fifth rows of Table 2. It can be seen, once again, that the stimulus uncertainties for the even-distribution sets of quotations are reflected quite faithfully, but that the uneven stimulus distributions are evened out. Nevertheless, a marked difference between even- and uneven-distribution sets is preserved by the response uncertainties.

### Comment

We find, therefore, that epistemic curiosity increases with evenness of distribution of supposed teachers' guesses and thus presumably with nearness to equality of strength of competing symbolic response-tendencies, now that extremely uneven distributions have been introduced. But this time, the number of alternative authors, which had such an impressive effect in Expt. I, fails to make any difference.

That one factor but not the other should exert a significant influence on the dependent variable is not in itself surprising, since rank-ordering data must, because of their limited information-transmitting capacity, show sensitivity for one independent variable at the expense of sensitivity for another. Nevertheless, the overall difference in response uncertainty between three-alternative and two-alternative quotations was actually slightly greater than the overall difference between even- and uneven-distribution quotations. It would seem that, if both number of alternatives and evenness of distribution affect epistemic curiosity, as the results of our two experiments suggest, these factors may not possess the same relative weightings in determining curiosity as they possess in determining uncertainty.

Part of this project was carried out during the tenure of an appointment as Visiting Scientist at the National Institute of Mental Health, United States Public Health Service. The remainder was supported by Research Grant M-4495 awarded by the United States Public Health Service.

The author is greatly indebted to the following who assisted in the collection of the data and the analysis of the results: Mrs Eunice Kennelly, Miss Geraldine Keene, Mr Paul Weisberg, Mrs Janet Barclay, Miss Nancy Leppert, Mrs Kathleen Spivack, Mr George H. Lawrence. He would also like to thank warmly the Principals of Bethesda-Chevy Chase High School, Maryland, and Lexington High School, Massachusetts, as well as their respective Boards of Education, for arrangements to collect data in their schools.

## REFERENCES

- BERLYNE, D. E. (1954*a*). A theory of human curiosity. *Brit. J. Psychol.* **45**, 180-91.
- BERLYNE, D. E. (1954*b*). An experimental study of human curiosity. *Brit. J. Psychol.* **45**, 256-65.
- BERLYNE, D. E. (1957). Uncertainty and conflict: a point of contact between information-theory and behavior-theory concepts. *Psychol. Rev.* **64**, 329-39.
- BERLYNE, D. E. (1960*a*). *Conflict, Arousal and Curiosity*. New York, Toronto and London: McGraw-Hill.
- BERLYNE, D. E. (1960*b*). Les équivalences psychologiques et les notions quantitatives. In D. E. Berlyne & J. Piaget, *Théorie du comportement et opérations*. (*Études d'Epistémologie Génétique*, XII.) Paris: Presses Universitaires de France.

(*Manuscript received 10 February 1961; revised 3 July 1961*)



## EXPERIMENTALLY INDUCED DRIVE AND DIFFICULTY LEVEL IN SERIAL ROTE LEARNING

BY R. A. WILLETT AND H. J. EYSENCK\*

*Institute of Psychiatry, University of London*

The interaction of experimentally induced drive with task difficulty was investigated in a group of candidates for selection to an industrial training school. The performance of the experimental group and a control group, on two nonsense syllable lists of different difficulty levels, showed no evidence of interaction but a simple facilitative effect of drive on level of performance.

The separate and combined influence of drive level and difficulty level upon both serial verbal learning and paired associate learning has received considerable attention in the recent literature. Most of this work has followed that of Spence and his collaborators (see review by Jones, 1960) in selecting as the drive condition under investigation the complex of secondary drives characterized as Manifest Anxiety and measured by the Manifest Anxiety Scale. There are certain objections to the measurement of the drive component by means of questionnaires, however, and a number of studies have attempted to avoid these by selecting or contriving their experimental situations to induce different drive levels (e.g. Beam, 1955; Sarason, 1956, 1957*a*, *b*). The interaction of these experimentally induced drives with such factors as difficulty level has then been examined. The present study is a contribution to this work and is part of a series of investigations into the role of drive in performance (Eysenck & Holland, 1960; Eysenck & Maxwell, 1961; Eysenck & Willett, 1961; Eysenck, Willett & Slater, 1962).

It should be noted that the emphasis upon the interaction of the drive component, with other variables like stress and difficulty level, in such studies as these, constitutes a departure from the Hullian position upon which they are based. It was hoped that the investigation being reported would throw some light on the relative merits of the interaction-type hypothesis and the basic Hullian concepts.

### SUBJECTS

A total of seventy-six young men were obtained through the apprentice training school of a large automobile manufacturing company. Seven records had to be rejected through subjects failing to complete the learning task in the time made available at the school. This point was reached after sixty trials had been completed.

### APPARATUS AND PROCEDURE

Each subject was required to learn a list of nine nonsense syllables to a criterion of two successive perfect repetitions. The procedure of serial anticipation was adopted and the method of presentation was by means of an electrically driven drum

\* We are indebted to Mr S. Rachman for assistance with the testing. The research was aided by a grant from the Society for the Study of Human Ecology and was made possible by the co-operation of Mr C. A. Attwood and the staff of the Apprentice Training School of the Ford Motor Company.

of the type employed by Hull (1940). Each syllable was presented for 2 sec., there being 4 sec. between the removal of the last syllable and the appearance of a cue to respond with the first syllable. The subject was given instruction in the method and required to learn a practice list of nine syllables to a criterion of two correct syllable anticipations. He was told to obtain the maximum score on each trial and to reach the criterion—two successive perfect repetitions—in the minimum number of trials.

#### EXPERIMENTAL DESIGN

Two sets of nonsense syllables were prepared from the lists published by Glaze (1928), one (the 'difficult' list) contained no syllable whose association value was more than 7%; the second (the 'easy' list) contained no syllable whose association value was less than 67%. Both lists conformed to the usual rules of constructing nonsense syllable lists.

Half of the subjects (thirty-eight) were taken from the 'Short List' (about eighty, selected from 600) of candidates undergoing selection to enter the apprentice training school. The other half were taken from students already undergoing training. Instructions were identical for both groups, but the former (the 'High Drive' group) carried out the learning task under examination conditions, as one of a number of tests comprising the selection battery. There was no evidence which suggested that these subjects did not believe, as was intended, the nonsense syllable learning task to be part of this battery. The group undergoing training (the 'Low Drive' group) were familiar with the experiments and with performance tests, although none had any experience of nonsense syllable learning. They were aware that they were being used as 'guinea-pigs', and, from past experience, knew that, although they were required to follow the instructions faithfully, their performance was unconnected with any reward and unrelated to their progress in the school.

Half of each group learnt the easy list and half the difficult list. Thus, there were four subgroups: (a) Low Drive, Easy list—'LE'; (b) Low Drive, Difficult List—'LD'; (c) High Drive, Easy List—'HE'; (d) High Drive, Difficult List—'HD'. After the rejection of subjects who were unable to complete the task the number of LE was 19, that of LD was 17, that of HE 18, and that of HD was 15.

Concerning the interaction between drive level and difficulty level two main predictions are possible. The prediction from classical Hullian theory would be simply that the high drive groups would learn more efficiently, no interaction being observed (and of course that the 'difficult' list would require more trials to the criterion). The prediction from hypotheses like Spence's and also from the Yerkes-Dodson Law would be that an interaction between drive level and difficulty level would be revealed, such that 'difficulty' would impair the performance of the high drive group relatively more than the performance of the low drive group.

#### RESULTS

Table 1 shows the standard deviations and mean number of trials taken by each subgroup to reach the criterion.

The subgroup arrays were subjected to an analysis of variance which involved the correction for disproportionality and unequal subclasses given by Snedecor (1956). The table of corrected sums of squares is given in Table 2.

Table 1. *Mean number of trials required to reach the criterion*

LE			LD			HE			HD		
No.	Mean	S.D.	No.	Mean	S.D.	No.	Mean	S.D.	No.	Mean	S.D.
19	28.84	11.69	17	40.71	9.54	18	25.22	10.23	15	33.07	10.79

Table 2. *Analysis of variance table. Number of trials required*

Source	D.F.	Sum of squares	Mean square	F	P
Between drive levels	1	515.06	515.06	4.33	< 0.05
Between difficulty levels	1	1694.83	1694.83	14.17	< 0.01
Interaction (drive $\times$ level)	1	69.90	69.90	1	N.S.
Error	65	7774.10	119.60		

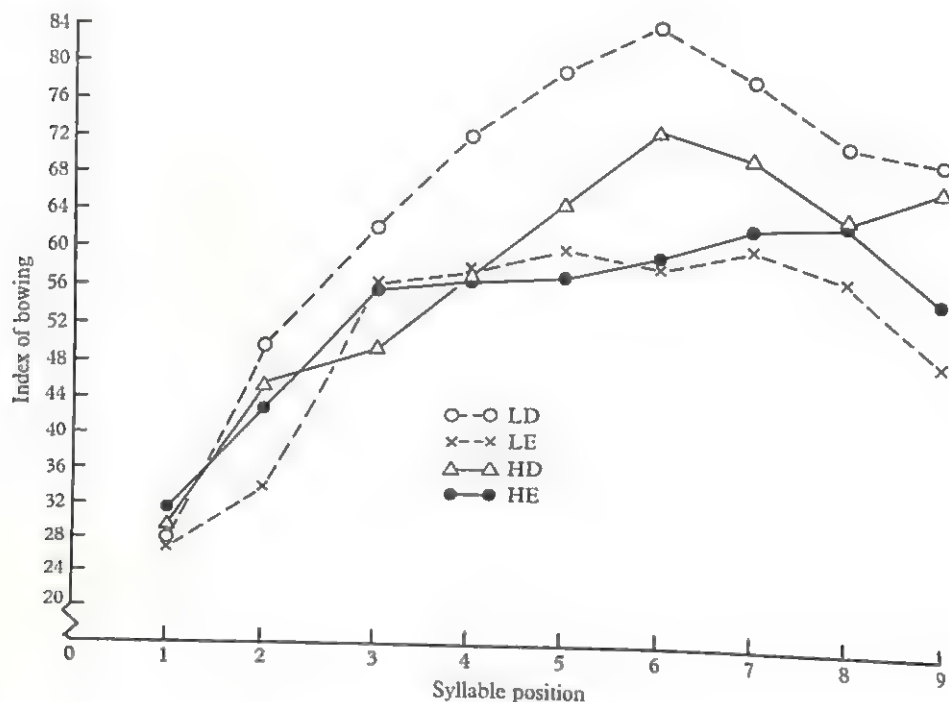


Fig. 1. Serial position effect.

It can be seen that superior performance was characteristic of the high drive group and of the 'easy' lists. No interaction between drive level and difficulty level was revealed.

Considerable speculation has centred on the distribution of errors across the item positions of rote learning tasks (the so-called 'bowed' serial position effect). However, whatever position is adopted the observation remains that the middle items of a list are relatively more difficult to learn than terminal items and we would thus expect either the greater efficiency of high drive in performance or its interaction with 'difficulty' to be particularly noticeable in the learning of these middle items relative to the terminal items.



Fig. 1. shows the distribution of errors across the items. These curves were constructed after a square root transformation had been applied to the error scores to equalize the variances of these scores across the item positions in order to obviate spurious differences in 'bowing' arising from the greater variances characteristic of the error scores of the middle items. After this transformation had been carried out, differences in 'bowing' were tested by obtaining, for each subgroup, an array of difference scores between the mean (transformed) error scores of items 1, 2, 3, 8, 9 and of items 4, 5, 6 and 7. The arrays of difference scores (multiplied by 100 to remove the decimals) so derived, constitute the indices of bowing. Table 3 shows the mean and standard deviations of these indices for the four sub-groups. All groups were reduced to fifteen to simplify computations in these calculations.

Table 3. *Indices of 'Bowing'*

LE			LD			HE			HD		
No.	Mean	S.D.	No.	Mean	S.D.	No.	Mean	S.D.	No.	Mean	S.D.
15	96.47	67.33	15	145.07	45.22	15	62.80	34.61	15	99.80	42.28

These subgroup scores were also subjected to an analysis of variance, the relevant table of sums of squares being reproduced below in Table 4.

It can be seen from Table 4 that greater bowing is associated with low drive and list difficulty. There is no evidence of any interaction between these two main effects.

Table 4. *Analysis of variance table—indices of bowing*

Source	D.F.	Sum of squares	Mean square	F	P
Between drive levels	1	23,364.26	23,364.26	9.08	< 0.01
Between difficulty levels	1	27,477.22	27,477.22	10.68	< 0.01
Interaction (drive $\times$ difficulty)	1	504.48	504.48	1	N.S.
Error	56	144,021.47	2571.81		

### DISCUSSION

The findings favour the classical Hullian account of the relationship between drive and performance as some simple progressive function and do not accord with those hypotheses which emphasize the interaction of drive with such variables as difficulty level. In both the case of over-all performance and of performance in the most difficult section of the task, the simpler Hullian model is adequate to account for the results without the necessity of postulating any interaction.

Quite clearly, however, these results alone cannot conclusively affirm the interaction-type hypotheses. The experimental design, here, can only arrange that different drive levels be induced in the two groups—it cannot ensure that the difference is such that an interaction with difficulty would necessarily be expected. If it were held that drive and difficulty interact at all levels, opportunity for them to do so was afforded in this investigation. If, however, it were argued that interaction occurs only within a narrow critical range of drive and difficulty values, then it is conceivable that the values obtaining in this investigation were inappropriate. To place

the matter beyond doubt it would be necessary to have some independent measure of the effects of the experimental conditions (possibly a P.G.R. record), to have a wider range of drive levels and lists even more 'difficult' than the 'difficult' list employed. If, under these conditions, interaction again failed to appear, the conclusion suggested by this investigation would receive considerable support.

#### SUMMARY

Two groups of subjects were formed such that the experimental situation induced a state of relatively high drive in one and of relatively low drive in the other. The 'high' drive group was selected from 'short' list candidates for entrance to the training school of a large industrial concern. The 'low' drive was selected from trainees attending the school familiar with experiments and performance tests. Half of each group was required to learn an 'easy' list and half a 'difficult' list of nonsense syllables. The interaction between drive level and difficulty level was examined and found to be negligible in both the case of performance to a criterion and over the middle items of the list. Superior performance and less marked 'bowing' of the serial position curve was associated with 'high' drive and the 'easy' list.

#### REFERENCES

- BEAM, J. C. (1955). Serial learning and conditioning under real life stress. *J. Abnorm. Soc. Psychol.* **51**, 543-51.
- EYSENCK, H. J. & HOLLAND, H. C. (1960). Length of spiral after-effects as a function of drive. *Percept. Mot. Skills*, **11**, 129-30.
- EYSENCK, H. J. & MAXWELL, A. E. (1961). Reminiscence as a function of drive. *Brit. J. Psychol.* **52**, 43-52.
- EYSENCK, H. J. & WILLETT, R. A. (1961). The measurement of motivation through the use of objective indices. *J. Ment. Sci.* **107**, 961-68.
- EYSENCK, H. J., WILLETT, R. A. & SLATER, P. (1962). Drive, direction of rotation and massing of practice as determinants of length of rotating spiral after-effects. *Amer. J. Psychol.* (to appear).
- GLAZE, J. A. (1928). The Association Value of nonsense syllables. *J. Genet. Psychol.* **35**, 255-67.
- HULL, C. L. (1940). *A Mathematico-Deductive Theory of Rote Learning*. New Haven: Yale University Press.
- JONES, H. GWYNNE (1960). Learning and abnormal behaviour. In H. J. Eysenck (Ed.), *Handbook of Abnormal Psychology*, pp. 488-528. London: Pitman Medical Publishers.
- SARASON, I. G. (1956). Effect of anxiety, motivational instructions and failure on serial learning. *J. Exp. Psychol.* **51**, 253-60.
- SARASON, I. G. (1957a). Effects of anxiety and two kinds of motivating instructions on verbal learning. *J. Abnorm. Soc. Psychol.* **54**, 166-71.
- SARASON, I. G. (1957b). The effect of anxiety and two kinds of failure on serial learning. *J. Personality*, **25**, 383-92.
- SNEDECOR, G. W. (1956). *Statistical Methods*. Ames, Iowa: Iowa State College Press.

(Manuscript received 20 March 1961)





## TOLERANCE FOR UNREALISTIC EXPERIENCES: A STUDY OF THE GENERALITY OF A COGNITIVE CONTROL\*

BY GEORGE S. KLEIN

*Research Center for Mental Health, New York University*

RILEY W. GARDNER AND HERBERT J. SCHLESINGER

*The Menninger Foundation, Topeka, Kansas*

A previous study demonstrated that subjects consistently differed in their willingness to experience perceptual organizations contradicting what they know to be true. Subsequent work has suggested that these individual differences reflect the operation of more general 'cognitive controls'. Results of the previous study, originally described in terms of 'tolerance' and 'intolerance' for one form of instability, are here conceived as representing a cognitive control more aptly described as 'tolerance for unrealistic experiences'. In an attempt to discover the range of applicability of this hypothesized control principle, and to gain further knowledge about it, 30 subjects were tested in a series of laboratory experiments. The Rorschach test was administered to 27 of these subjects. The tests were chosen to bring about unconfirmable, i.e. unrealistic, experiences by two general means: (a) by manipulating the physical stimulation so that reality is altered; and (b) by presenting subjects with relatively ambiguous stimuli for which 'reality' provides no clear-cut confirmation, so that subjects are free to organize them in a number of ways.

The results generally support those of the original study and extend the range of situations in which the control principle can account for individual differences.

### I. INTRODUCTION

A previous study (Klein & Schlesinger, 1951) demonstrated that subjects who showed attitudes of tolerance and intolerance for instability in their Rorschach responses differed in their readiness to experience apparent movement. In that experiment, 'instability' was created by evoking a discrepancy between what is known and what is perceived. 'Tolerance for instability'—here referred to as 'Tolerance for unrealistic experiences'—was indicated by the subject's readiness to accept and report experiences at variance with conventional reality or with what they knew to be true.

To establish a bridge between performance on the Rorschach and on the apparent-movement test, it is necessary to understand the adaptive problem confronting the subject in both. The subject is fully aware that the stimuli producing apparent movement are actually stationary (in our studies this has been explained and demonstrated to him). Under these conditions the range of alternation rates over which he reports movement indicates the degree of his readiness to compromise with the known state of affairs.

A Rorschach inkblot is also unstable. The adaptive task defined by the experimenter is open-ended; he offers no rules of interpretation to guide the subject in moulding his responses or in choosing among equally plausible alternatives. If the instructions are taken literally, any communicable response is acceptable. The subject must decide how far he will depart from the 'known', most certain, least contestable,

\* These experiments formed part of a research programme conducted under grants from the National Institute of Health, United States Public Health Service.

identity of the stimulus (the fact that it is an inkblot) and allow himself to experience the blot's myriad organizational possibilities.

Thus, the apparent-movement and Rorschach situations bring about non-confirmable experiences in different ways. In the apparent-movement test, perceptions consonant with what the subject accepts as 'real' are made increasingly untenable by physical variations in the conditions of stimulation, i.e., by increasingly rapid alternations. The physical changes coerce experience away from one that is in harmony with environmental fact, towards one that is not. In the Rorschach, the subject's tolerance for, and modes of dealing with, equivocality are put to the test by the many equally legitimate response possibilities, no one of them more 'correct' than another (although, to be sure, 'popular' responses imply that there is relatively more unequivocal stimulus support for certain possibilities than others), and by the mild pressure, implicit in the instruction 'What might this be?', to perceive more than the literal givens of the blots.

The present report further explores the usefulness of this control principle in a variety of tasks which pitted different experiential possibilities against each other. The test situations involved apparent movement; reversible figures; distortion of the visual field induced by aniseikonic lenses; and the Rorschach test (applying criteria derived from an earlier study and improved upon). In different ways these situations produced equivocation and even contradiction among the informational leads made available to a subject.

In the previous study, the Rorschach test was used as the criterion situation. In the present series, the range of apparent movement was the criterion measure, i.e., the number of cycles per second between the transition point from alternation to movement and the transition point from movement to simultaneity. The strategy was to predict responses to each of the other tasks from apparent movement scores. Because the score distributions in the three experimental tasks varied in shape, all measures except those used to determine consistency of apparent movement performance were T-scaled for the correlational analyses.

## II. SUBJECTS

Subjects were 30 patients and employees (largely clerical workers) of a midwestern psychiatric hospital. Thirteen were men, seventeen women. Their ages ranged from 19 to 43, with a median of 31.0. Psychiatric diagnosis was ignored in the selection of the subjects. Chi-squared tests were not significant for any of the experimental variables when subjects were divided on the basis of sex, age, or patient versus employee status, so that these were ignored in analyses of the data.\*

\* Whether or not extremes of the control principle discussed here are associated with particular neurotic syndromes is a matter for further research. In respect to the mean range of apparent movement, however, these results are compatible with Hamilton's (1960) finding that control and neurotic groups differ significantly in movement *interval* (periods of repeated movement), but not in movement *range*.



## III. TOLERANCE FOR UNREALISTIC EXPERIENCES IN AN APPARENT-MOVEMENT TEST

In the earlier study (Klein & Schlesinger, 1951) members of the 'intolerant' (T-) group preferred perceptual organizations congruent with conventional reality. In the apparent-movement test, these subjects were governed by their knowledge that *stationary* figures were being presented, even when alternation was well within the range of readily perceivable movement.\* Members of the 'tolerant' (T+) group were less constrained by this knowledge of the stimulus conditions; they were more able to accept the movement experience through a broader range of alternation rates.

(i) *Apparatus and procedure*

The apparatus consisted of a Dodge-type tachistoscope modified to provide three stimulus fields. The stimuli were drawn in black India ink on 8 × 5 in. white file cards and were approximately  $\frac{1}{4}$  sq. in. in area. The pairs of stimuli, similar to three of the six pairs used in the earlier study, were (a) running horses; (b) identical squares; (c) bars arranged |—. Fields *A* and *B* held the two stimulus cards, field *C* a blank card. The fields were illuminated at a comfortable viewing level by 'daylight' fluorescent tubes. The stimuli were exposed in *ACBC* order. Intervals between appearances of the stimuli were filled by a white background of the same intensity as the two stimulus cards. Flicker was thus reduced to a minimum. The subject viewed the stimuli, 27 in. away, through the eye-piece of the tachistoscope. Alternation rate was controlled by means of a motor-driven commutator, variable in speed from 0 to 8 revolutions (cycles) per second. The mechanism consisted of a  $\frac{1}{8}$  h.p. motor that drove a dual-controlled, variable-speed transmission. To eliminate noise, these parts of the apparatus were isolated in another room and controlled by means of a flexible shaft running through the wall.

Before the test trials, the subject was shown the apparatus, including the stimuli, and the means of producing the illusion of movement was demonstrated to him. Enough demonstration trials were given to insure that he could recognize the points he would be asked to report in the test trials (i.e. alternation, movement, simultaneity).

In the five test trials with each of the three pairs of stimuli, the experimenter gradually stepped up the alternation rate until the subject reported that the two alternating figures now moved as a single figure. (In the earlier experiment, the subject had controlled the alternation rate.) The experimenter then continued to increase the alternation rate until the subject reported two simultaneously flickering figures. Without stopping, the experimenter increased the alternation rate further, until he was confident that the flicker experience was a stable one. Subjects occasionally reported both flicker and movement within a brief range of alternation rates. In such instances, the alternation rate was increased to the point where flicker became the dominant and stable experience, and this was recorded as the upper limit of the subject's range of apparent movement.

\* The influence of central cognitive factors upon apparent movement is indicated by Jones and Bruner's (1954) demonstrations of the effect of expectancy, and by Toch & Ittelson's (1956) evidence of the impact of past experience.



The datum for each trial was the range of alternation rates between the first point (movement experience) and second point (flicker experience), in cycles per second. The measure for each stimulus figure was the subject's mean range of apparent movement for the five trials. Subject consistency was probed by obtaining Pearson  $r$ 's between raw-score mean ranges for the three pairs of stimuli. The  $r$ 's were (a) horses: squares, 0.90,  $P < 0.001$ ; (b) horses: bars, 0.72,  $P < 0.001$ ; (c) squares: bars, 0.76,  $P < 0.001$ . The criterion measure of tolerance for unrealistic experiences was the mean range of apparent movement of the fifteen trials with the three pairs of figures. The scores were then T-scaled for purposes of prediction to performances in Expt. 1-3.

#### IV. EXPERIMENT 1. TOLERANCE FOR UNREALISTIC EXPERIENCES IN REVERSIBLE-FIGURE TESTS

The typical reversible figure can be seen in either of two phases. With steady viewing, these phases alternate. Although the subject knows that the change is apparent, not real, the phenomenon is unmistakable and compelling. While the mechanism of alternation is not clearly understood, it is likely that central organizing factors have some part in determining the wide range of differences among people in the rate of reversals that have been noted by many experimenters.

Reversible figures seem to exhibit the qualities of nonconfirmability referred to earlier. Instability is inherent in these figures in that they produce unequivocal experiences, which invite resolution in the sense described for the apparent-movement test, except that the experimenter does not manipulate the tenability of an experience by varying the stimulation. Also, neither one of the two possible organizations is more 'correct' than the other, and in this respect reversible-figure tests exhibit the quality of equivocality characteristic of the Rorschach test.

*Hypothesis 1.* There is a positive relationship between mean range of apparent movement and number of reversals, both when subjects are asked to attend *passively* to reversible figures and when they are asked to *resist* experiencing alternation.

*Hypothesis 2.* There is a negative relationship between mean range of apparent movement and the time during which that phase of a reversible figure is experienced which is in greater agreement with conventional, everyday experience.

##### (i) *Apparatus and procedure*

A 2 × 4 in. Schroeder 'staircase' provided a figure of contrasting conventionality in its two phases. In one phase, this figure is a conventional staircase. In the other, it conforms less to everyday experience. The second figure was a 'windmill' three inches in diameter, with four black and four white vanes. In this figure, the two phases are obviously less different in their conventionality than those of the 'staircase', although pilot experimentation indicated that for most subjects the black 'windmill' is the dominant organization. These figures were drawn in India ink on 5 × 8 in. white file cards. The subject viewed them one at a time at a distance of three feet. The stimuli appeared against a black cardboard background and were viewed under normal room illumination.

Before the experimental trials, the phenomenon of reversal was described to the subject and he was allowed to examine the figures and to experience them in both

phases. During the experimental trials, he held a telegraph key that was connected to the clutch of an electric stoplock. He depressed the key whenever the 'staircase' appeared inverted and kept it depressed until the figure 'reversed'. For the 'windmill', the key was depressed when the figure was seen as a white windmill on a black background.

There were two 1 min. trials with each figure in the 'passive' condition, in which the subject simply viewed the figure steadily and indicated reversals by depressing or raising the telegraph key. The trials were administered in *ABBA* order. In a subsequent 1 min. trial with each figure (*BA* order), he was asked to 'hold back' alternation by 'pitting his will' against it. The instructions emphasized that neither position of the figure was to be deliberately favoured and that no artificial means of reducing reversals, such as squinting, were to be used. Scores were (a) number of reversals, and (b) the conventional-phase time for each figure in each condition. The scores were T-scaled.

### (ii) Results

The results were in accord with Hypothesis 1 except for the windmill figure in the 'passive' condition. The  $r$ 's between mean range of apparent movement and number of reversals were: staircase figure, 'passive' condition, 0.53,  $P < 0.005$ ; staircase figure, 'hold-back' condition, 0.47,  $P < 0.005$ ; windmill figure, 'passive' condition, 0.05; windmill figure, 'holdback' condition, 0.40,  $P < 0.025$ . Possibly the near-equivalence of the 'stand-out' figure in the two phases in respect to conventionality complicated the results for the windmill stimulus. A stronger discrepancy in the relative conventionality of the alternate organizations seems to be necessary to bring this dimension into play. The inhibitive effort of the 'hold-back' condition perhaps tended to highlight the otherwise small discrepancy between the two figures in conventionality.

Hypothesis 2 was not confirmed, except for the 'hold-back' condition with the windmill figure. The  $r$  between mean range of apparent movement and conventional-phase time was 0.037,  $P < 0.025$ . For the 'passive' condition with the windmill figure,  $r$  was -0.17. For the 'passive' and 'hold-back' conditions with the staircase figures, the  $r$ 's were -0.05 and -0.09.

The results, in general, suggest that perception of reversals reflects tolerance for unrealistic experiences in the sense described, particularly when there is a marked difference in the conventional 'reality' of the two phases of the figure.

## V. EXPERIMENT 2. TOLERANCE FOR UNREALISTIC EXPERIENCES INDUCED BY ANISEIKONIC LENSES

Spectacles fitted with aniseikonic lenses (described by Ames, 1946) create a disparity in the sizes of the two retinal images that is greater than the normal disparity produced by the distance between the pupils. In one position of the 2% meridional lenses used in the present experiment, axes are positioned obliquely 45° right eye and 135° left eye. Besides changing the apparent sizes and shapes of objects, this type of aniseikonic lenses induces a decided tilt in the vertical-frontal plane. In the position referred to, the lenses make the floor slant sharply downward and the top of the wall in front tilt away. The experienced tilt can be reversed by inverting the lenses. Thus,



the unrealistic experience induced by aniseikonic lenses, like that in our apparent-movement test, is brought about by an alteration of proximal cues.

Ames (1946) has pointed out that the effects of aniseikonic lenses vary with environmental conditions: (a) the availability of 'unioocular stimulus-patterns with which visual significances are related' (p. 354); (b) the distance at which objects are viewed; (c) the presence of familiar objects whose ordinary spatial position is well known to the subject.

The effects also vary among individuals. Ames suggested that subjects' relative sensitivity to unioocular and binocular cues may be a major source of these differences. Possibly, the T+ and T- attitudes may affect this sensitivity in the aniseikonic lens situation. Thus, T- subjects would be expected to maintain the customary rectilinearity of the laboratory by attending to the unioocular cues that help maintain it. The possible relevance of the aniseikonic lenses experience to this dimension was suggested by Kaplan's finding (1952) that lens scores, similar to those of the present study, correlated significantly with facility in recalling contradictory and loosely related-story elements, and by Gardner *et al.* (1959). Martin (1954) has also reported significant relationships between scores roughly similar to those used in the present experiment and the number of questions subjects asked to clarify three ambiguous interpersonal situations.

*Hypothesis 1.* There is a negative relationship between mean range of apparent movement and time required to 'recognize' the tilt induced by the lenses.

*Hypothesis 2.* There is a positive relationship between mean range of apparent movement and the amount of tilt experienced.

#### (i) Apparatus and procedure

The subject sat in a desk-chair facing a wall covered with a white cloth stretched taut so as to present a uniformly bright surface. The desk portion of the chair was also covered with a white cloth. Floor, ceiling, and adjacent wall (painted concrete blocks) were visible to the subject when he moved his head.

Amount of distortion experienced by the subject with the lenses on was measured by having him adjust a black wooden rod to the vertical. The rod,  $\frac{3}{4}$  in. square by 15 in. long, with rounded ends, was mounted at the far edge of the desk-chair on a friction hinge, so that the subject could adjust its angle of inclination toward and away from him. Deviations from the vertical were measured with a bubble protractor.

Prior to wearing the lenses, the subject was asked to make five adjustments of the rod to the vertical. No one of these adjustments by any subject varied more than a few degrees from the vertical. The average error of the group was near zero.

During the first part of the test, the lenses were worn in the position (axes 45° right eye, 135° left eye) that made the floor seem to tilt downward in front (position 1). With eyes closed, the subject put on the spectacles and was told that 'things seem quite different to some people when they look through these glasses, while for others this is not the case'. He was asked to report 'how things look' after he opened his eyes and to describe in detail any differences from the usual appearance of the room. He was encouraged to look around the room in a normal manner while wearing the lenses. The experimenter then asked him to open his eyes, recorded his comments,



and noted the time when he first commented upon tilt. The subject's experience was then reviewed with him as a check on the recorded recognition time. If he did not report experiencing distortion in 10 min., he was queried to determine whether or not he had experienced distortion but had failed to report it. It was decided beforehand that only subjects who had reported distortion sometime during the allotted 10 min. would be included in analyses of relationships of this task with others. Following this, five measurements of experienced distortion were made by having the subject adjust the rod to the 'true vertical'.\*

Scores were the time in seconds required to report tilt initially (recognition time), and mean deviation from the true vertical. The scores were T-scaled.

### (ii) Results

Hypothesis 1 was supported by the  $r$  of  $-0.39$ ,  $P < 0.025$ , between mean range of apparent movement and recognition time. Although the  $r$  between mean range of apparent movement and mean *degrees* distortion was in the expected direction ( $0.16$ ) it did not reach significance.

### (iii) Discussion

The present results are consistent also with those of Snyder (1956), who found relationships between responses while wearing aniseikonic lenses, the Kohnstamm phenomenon, and a battery of Rorschach measures similar to those used by Klein & Schlesinger (1951). It must be noted, however, that the critical correlation of  $-0.39$  in the present study of the hypothesized relevance of the lens experience to T+ and T- controls accounts for only 15% of the variance. Very likely determinants other than those proposed by Ames (unocular and binocular clues) and possibly other dimensions of cognitive control contribute to experiences of tilt and to recognition delay. The latter possibility is underscored by the results of the factor analysis reported later in this paper.

Previous studies of cognitive control principles have taken for granted their automatic, non-volitional character; they guide behaviour without the individual's explicit awareness (Gardner *et al.* 1959). An indication of this appeared in the curious finding that two subjects denied seeing change in the visual field while wearing the aniseikonic lenses yet showed considerable deviation in their adjustments of the rod to the vertical. One of these subjects was near the 'intolerant' extreme of the distribution of ranges of apparent movement in this sample, the other near the most 'tolerant' in terms of this criterion measure. Evidently degree of awareness of the changes induced by the lenses is not correlated in any simple fashion with tolerance for unrealistic experiences.

\* Measurements of the experienced tilt with the lenses in the *inverted* position (axes  $135^\circ$  right eye,  $45^\circ$  left eye) are confounded by an artifact, and therefore the results for this position are not reported here. With the lenses in this position, adjustment of the rod in conformance to the apparent tilt of the wall brought the top of the rod close to the subject's face. All of the subjects showed less deviation of the rod from the vertical with the glasses in this position, and the range of deviations was smaller than for position 1. Thus, position 1 seemed to provide a more adequate measure of perceived distortion.

# VI. EXPERIMENT 3. TOLERANCE FOR UNREALISTIC EXPERIENCES IN THE RORSCHACH TEST

Behaviour in the Rorschach test and the experiencing of apparent movement were linked in an earlier study (Klein & Schlesinger, 1951). What we have referred to here as intolerance for unrealistic experiences manifested itself in that study in a diffidence to take up the challenge offered by the 'open-endedness' of the Rorschach test. The main qualitative attributes of T- protocols were (a) concern over the reasonableness of responses; (b) literalness of approach, i.e. a tendency to anchor responses to unequivocal attributes of the stimuli; (c) a tendency to report only clearly delineated, easily confirmable forms and meanings; (d) avoidance of associative elaboration. Contrasting properties characterized the T+ protocols.

The general hypothesis in the present experiment was that there would be agreement between the T- and T+ criteria of apparent movement range and the T- and T+ classification of subjects on the Rorschach test. Thus, subjects with large mean ranges should be classified T+ on the Rorschach, subjects with small mean ranges should be classified T-.

## (i) *Rorschach criteria for judging tolerance and intolerance for unrealistic experiences*

The rating criteria were modified and elaborated on the basis of the earlier study, and were applied by experienced clinicians to the protocols of the present subjects. The criteria were developed on the premise that the T+ and T- extremes are more readily discernible in a subject's attitude toward his responses than in the standard 'signs' by which Rorschach protocols are usually scored.

### (a) *Evidences of concern with realism and reasonableness of responses*

The main difference between T+ and T- is the freedom the subject permits himself of tampering with the 'reality' of the card. Differences in this respect are revealed in a number of ways. The T+ person comfortably accepts the task as an opportunity for projection. He may even view the blot as something to be played with. If he does not necessarily enjoy the task, he at least does not find it uncomfortable. A T+ protocol is notably free of critical comments and expressions of dissatisfaction with the task. The T+ subject can toy with alternative conceptions for a given area; he may elaborate responses from small hints provided by the blots. In formulating meanings for the blots he may draw either upon the train of his associations, or on specifiable physical areas; he does not feel compelled to anchor responses only to meanings that are clearly confirmable by physical features of an area nor only to physical details that provoke least doubt about the meanings they suggest.

The T- person, on the other hand, sets limits to the ideational freedom he permits himself in confronting the open-ended instruction. This shows itself not so much in unproductivity or few responses as in a tendency to subject his responses to a critical eye (e.g. 'A bat, but not a very good one'). He prefers certainty, the confirmable meaning; the less certain is allowed expression only grudgingly, warily, and with circumspection. His protocols are dotted with qualifiers as if 'to keep the



record straight' (e.g. 'It looks a little bit like...'). He is concerned over an imaginary picture that does *not* fit a meaning; he tends generally to be caught up in the formal qualities of the *blot* rather than in the associative byways of the meanings that happen to come to mind.

The T- orientation may disclose itself less in these specific ways than in a general discomfort with the task itself, indicated by the subject's manner of response, e.g. he may quibble; he may repeatedly ask questions about how he is to respond; he may complain about the vagueness of what is expected of him. This discomfort may take the form of projected meanings (e.g., a person or animal may be 'grotesque', a form 'vague'). Although he may give many responses, the quantity will not testify to ease and freedom of responsiveness.

(b) *Variety of determinants:*

T+ qualities may be revealed through a variety of other determinants than form alone, such as shading, colour, etc., all, for the most part, integrated with the form-attributes of the blots. References to colour, for example, should disclose this easy blend. In contrast, the T- attitude may be implemented either by a total unresponsiveness to colour or by sporadic, unmodulated reactions to it. Though by itself it is not as reliably indicative of the T- extreme as other criteria, such 'all-or-nothing' responsiveness to colour did tend to appear more frequently in the earlier study in the context of T- protocols.

It was possible to recall for Rorschach testing all but three of the 30 subjects who had participated in Expt. 1 and 2. The missing three were subjects previously classified as T+ on the apparent movement test. Therefore, the raters were asked to sort fifteen of the records as relatively T- and twelve as relative T+, to conform to the numbers so classified by the apparent movement criterion.

The raters, three experienced clinical psychologists, read each Rorschach protocol in its entirety and classified it as relatively T- or T+ on the basis of the qualitative characteristics outlined. The raters worked independently, and a two out of three consensus determined the final subject groupings.

Agreement among the raters in making this two-category sort was assessed by  $\chi^2$  tests. Agreement was satisfactory: Raters A and B,  $P < 0.025$ ; A and C,  $P < 0.10$ ; B and C,  $P < 0.025$ .

(ii) *Results*

The biserial  $r$  between the T+ and T- groups formed by the Rorschach sorting and the criterion score (mean range of apparent movement) was 0.65,  $P < 0.001$ . In confirmation of the earlier study, tolerance for unrealistic experiences is reflected both in the Rorschach test, as standardly administered, and in apparent movement responses.

The Rorschach protocols were further reviewed by the experimenters, who summarized major qualitative attributes of subjects who fell into the T- or T+ halves of the apparent movement scores. Again, the Rorschach records of the two groups displayed qualities similar to those observed in the first study. The following sample comments were extracted from the summaries of protocols of T- apparent movement subjects.



'He is continually aware of the quality of his responses and continuously evaluates their closeness to reality'; 'he emphasizes self restraint and control'; and 'he attempts to justify responses by enumerating details and congruences'.

'He displays little conviction in his responses; spontaneously justifies his responses by finding details, offering to show the experimenter where the response is.'

'He is anything but relaxed in giving his responses.'

'She lacks any basic conviction in what she sees... is uncomfortable in the test.'

Three members of the T- group gave large numbers of responses, but these were amply accompanied by indications of the T- attitude described in the rating criteria. High *R* is evidently the outcome of a quite independent need to produce. In the T- context, the high productivity seemed driven, rather than free, a quality well illustrated in a subject with 96 responses to the blots. The records were distinguished by barrenness of content, lack of confidence in the adequacy of responses, over-concern for the 'real' characteristics of the blot, a need to justify responses and to seek explicit stimulus confirmation. It was evident that such productive T- people organized the task in the letter rather than the spirit of phantasy; in a variety of ways, they underscored the shaky commitment they felt for their responses.

The T+ records were qualitatively less homogeneous than those of the T- group, although generally consistent with the T+ tendencies noted in the previous study. Typical interpretive remarks in the qualitative summaries of these records were:

'Initially insecure, but before long warms up to the task and gets free enough to "play" associatively with the cards... never loses all concern with reality, but can joke about it.'

'... there seems not to be much concern for reality limitations; fluid... confabulates without question.'

Table 1. *Intercorrelations of major scores\**

Variable	No.	(T-scaled data.)									
		Correlation									
Apparent movement, mean range†	1	—	—	—	—	—	—	—	—	—	—
Staircase, reversals, passive†	2	53	—	—	—	—	—	—	—	—	—
Staircase, reversals, hold back†	3	47	55	—	—	—	—	—	—	—	—
Windmill, reversals, passive	4	05	24	14	—	—	—	—	—	—	—
Windmill, reversals, hold back	5	40	57	61	40	—	—	—	—	—	—
Staircase, conventional-phase time, passive†	6	-05	-39	01	07	-23	—	—	—	—	—
Staircase, conventional-phase time, hold back†	7	-09	03	-16	-05	-05	24	—	—	—	—
Windmill, conventional-phase time, passive	8	-17	-21	-26	-22	-36	36	20	—	—	—
Windmill, conventional-phase time, hold back	9	-37	-41	-50	-04	-31	24	37	59	—	—
Aniseikonic lenses, recognition time†	10	-39	-30	-14	-28	-22	14	-09	08	11	—
Aniseikonic lenses, mean deviation 1†	11	16	00	-15	24	-12	13	08	22	-05	-52
Rorschach classification ††	12	65	06	16	08	31	15	-31	-30	-58	-12
		1	2	3	4	5	6	7	8	9	10

\* Decimal points omitted. † Used in the factor analysis. ‡ Biserial *r*'s.

## VII. FACTOR ANALYSIS

(i) *Statistical procedure*

In an attempt to assess the over-all importance of tolerance for unrealistic experiences as a major determinant of scores in the three experiments, a factor analysis was carried out by Thurstone's (1947) centroid method. The complete correlational matrix appears in Table 1. In order to avoid undue distortion of the factor structure, only eight relatively unique scores were included in the factor analysis.

The maximum number of significant factors was three as determined by three tests of significance [Tucker's phi, Humphrey's rule, McNemar's criterion, as described by Fruchter (1954)]. These factors were rotated orthogonally (a) graphically and (b) analytically, by the normal varimax method (Kaiser, 1958).\*

(ii) *Results*

Results of the graphic rotation appear in Table 2, results of the analytical solution in Table 3. The two methods yielded similar results. By means of hand rotations it was possible, without deviating from the criterion of orthogonality, to achieve defining loadings† on Factor I for five of the scores presumed to represent tolerance for unrealistic experiences, as against four loadings produced in the analytical solution. The interpretations below are based on results of the graphic rotations.

Table 2. *Rotated orthogonal factor loadings (graphic)\**

Variable	Factors			$h^2$
	I	II	III	
1	76	-35	-08	71
2	35	-16	-66	58
3	53	11	-34	41
6	-01	00	58	34
7	-37	-19	14	19
10	-05	77	21	64
11	06	-64	13	43
12	74	-27	33	73
Percentage of variance accounted for	21	16	14	

\* Decimal points omitted.

(iii) *Interpretations of the factors*

*Factor I.* This factor, accounting for a greater percentage of score variance than either of the other two factors, is clearly interpretable in terms of tolerance for experiences that are at variance with conventional reality. Except for failure of the aniseikonic lens scores to show defining loadings, the factor corresponds well to the assumptions presented earlier which guided the interpretation of individual scores in terms of this principle of cognitive control. It will be noted that the conventional-

\* We are indebted to Dr Steven G. Vandenberg, of the University of Michigan, for providing us with an electronic computer programme written by Dr Donald Lamphiear for the normal varimax method, and to Dr Harriet B. Linton, of the Research Center for Mental Health, New York University, for performing the graphic rotations.

† Factor loadings of 0.35 and above were accepted as defining loadings in the present study.

phase time score for the staircase figure in the hold-back condition has a negative loading, as predicted. The fewer number of reversals in T- subjects may have been brought about by a preference for the more conventional upright phase of the figure.

Table 3. *Rotated orthogonal factor loadings (normal varimax)\**

Variable	Factors			$h^2$
	I	II	III	
1	75	-35	16	71
2	28	-15	70	59
3	50	12	39	42
6	05	-01	-58	34
7	-35	-19	-17	19
10	-07	76	-24	64
11	07	-64	-12	43
12	76	-28	-25	72
Percentage of variance accounted for	20	16	15	

\* Decimal points omitted. Factors numbered according to the percentages of variance accounted for.

*Factor II.* This factor may represent another cognitive control principle, levelling-sharpening (Holzman, 1954; Gardner *et al.* 1959), which was proposed to account for individual differences in assimilation-effects. The high loadings all involve scores that could represent the degree to which distinctive features of stimuli are elided through the assimilation of stimulus input to a 'normative' level established by previous experience. For instance, while the aniseikonic lenses produce changes in proximal cues, these can still be encompassed by the strongly established, well-confirmed schema of the room as rectangular. It is, therefore, the 'norm' rather than the deviant proximal cue that dominates report. The schema determined the experience 'without correction', as it were. The degree of such assimilation may thus contribute both to the delay in recognizing tilt and to the degree of tilt experienced.

A similar assimilation tendency may have affected apparent movement experience under the conditions of the present study. As the test was given, the initial dominant experience was alternation. Where a subject's assimilation tendency is pronounced, the gradually more rapid alternation may be referred to this 'norm', thereby prolonging the 'alternation' phase of response, with a consequent reduction of the range of experienced movement. In such a subject it takes a larger-than-usual increment of stimulation (increasing alternation rate) to make the alternation experience untenable.

Only orthogonal rotations were used, since they require a minimum of assumptions. However, the plot of Factor I against Factor II clearly indicated that an oblique rotation would best fit the data. The appropriate oblique rotation of Factors I and II yielded two factors similar to the present ones but correlated 0.45. A common element, presumed to be tolerance for unrealistic experience, is apparently involved in the two factors. The discreteness of the factors perhaps arises from the assimilation-tendency component of Factor II. Had more direct measures of levelling-sharpening (Holzman, 1954) been included in the present study, Factor II might have pulled apart more definitely from Factor I.

*Factor III.* This factor seems to represent mainly responsiveness to the staircase



figure (the near-defining loading for the Rorschach classification is puzzling in this context). Since the factor is orthogonal to Factor I, it represents still another determinant of the dominance of the upright phase of the staircase figure in the passive condition, in addition to that of relative conventionality of meaning which we have assumed distinguishes the two organizations. The interpretation of this factor is uncertain.

### VIII. DISCUSSION

The present study suggests that tolerance for unrealistic experiences can be observed when experiences of an unstable stimulus field differ from each other in respect to their congruence with common experience. In some of the tests, the experimenter violated the subject's conviction of what was true by physically manipulating the stimuli to bring about a contradictory experience (e.g. apparent motion). Other tests (e.g. Rorschach) allowed subjects a variety of response options through which to express their readiness to experience unconventional, less readily confirmable, organizations. The factor analysis supported the interpretation that tolerance for unreality, so defined, contributed to performances in the tests used, although it disclosed other possible sources of individual performance. For example, assimilation effects may be involved in the aniseikonic-lenses tests used, raising the possibility that another cognitive control principle, levelling-sharpening (Holzman, 1954), may affect recognition time for tilt and the degree of tilt experienced.

The results of the present study provide a bridge to Kaplan's (1952) finding that acceptance of lens-induced tilt is significantly correlated with recall of two kinds of 'unstable' (i.e. contradictory and loosely related) story elements. Accessibility of material for later recall may be affected by the subject's manner of dealing with combinations that defy logic or realism. Thus, it is not 'reasonable' that a single story should contain elements that directly contradict each other; accordingly, T- subjects respond by not reproducing them. The fact that T- subjects also recalled fewer *loosely related* story elements suggests that the principle may have still more general attributes. However, Kaplan used the aniseikonic-lenses test to provide criterion measures, and a further experiment is needed to determine the contributions both of tolerance for unrealistic experiences and of levelling-sharpening (Holzman, 1954) to the response characteristics he observed.

Martin's (1954) finding that individual differences in response to ambiguous social situations can be predicted from experiences while wearing aniseikonic lenses also suggests a possibly fruitful extension of the tolerance for unrealistic experiences dimension. However, the relationship between this principle and responsiveness to ambiguity *per se* (cf. Frenkel-Brunswik, 1949; Hamilton, 1957; Loomis & Moskowitz, 1958) needs to be explored.

While, in general, the various situations sampled in the present study seem relevant to the hypothesized control principle it may be helpful for future, more systematic exploration of this dimension to indicate some general classes of task conditions which may be specially conducive to the appearance to T+ and T- controls:

(1) T+ and T- will appear under conditions that create an explicit discrepancy between informational input, on the one hand, that bolsters a belief in what is 'real', and, on the other hand, contiguous informational input that produces an experience

contradictory to this belief (e.g., in the apparent motion test: discrepancy between the experimenter's assurance that the lights are always stationary and the motion experience induced by rapid alternation).

(2) T+ and T- attitudes will appear under conditions where the subject carries into the situation a strongly habituated, long practiced and verified conviction which is now countermanded by a systematic and stable change of proximal stimulation. The aniseikonic lenses situation is of this character. Tasks of this type would constitute a problem in the relative sturdiness of well-established adaptation levels.

(3) T+ and T- attitudes will appear under conditions where a perceived informational return (feedback) from an action or a perception turns out to be misleading, i.e. produces false hypotheses of ineffectual actions (delayed feedback, mirror-writing, etc.).

(4) T+ and T- attitudes will appear under conditions which foster rapid changes in the contents of awareness.

The tests in the present study are all subsumable under conditions (1) and (2). In this connexion, it would be interesting to complement the apparent movement test of the present study with another in which subjects are led to believe that the stimuli are in fact *moving*. We have seen that apparent movement range takes on meaning in terms of this pre-knowledge. It is our plan now to tell subjects that the stimuli are actually moving (for demonstration, an optimal alternation rate is selected which is seen by all as movement). As before, we will proceed to reduce alternation rate and determine the critical point of a report of alternation. We expect that T- subjects who in the original test gave a narrow range will now give a broad range; the reverse should be true for T+ subjects.

Conditions (3) and (4) are not represented in the present study, but suggest likely conditions for investigating tolerance and intolerance for unrealistic experiences. Condition (3) presents a situation in which the reliably informative feedback on which people ordinarily depend in guiding their behaviour perceptually is no longer forthcoming. The character of the behavioural disruption and of readjustive efforts should differ in T+ and T- subjects.

Condition (4) proposes that the T+ and T- controls may be reflected in differences among people in their inclination to recall or dwell upon transitory, unstable, fluid hypnogogic and hypnopompic imagery, assuming they even have such experiences. Similarly, the organization of recalled dreams in T+ and T- subjects may differ. These possibilities will be explored in future studies.

#### REFERENCES

- AMES, A., Jr. (1946). Binocular vision as affected by relations between uniocular stimulus-patterns in commonplace environments, *Amer. J. Psychol.* **59**, 333-57.
- FRENKEL-BRUNSWIK, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Personality*, **18**, 108-43.
- FRUCHTER, B. (1954). *Introduction to Factor Analysis*. New York: Van Nostrand.
- GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, H. B. & SPENCE, D. P. (1959). Cognitive control: a study of individual consistencies in cognitive behavior. *Psychol. Issues*, **1**, no. 4.
- HAMILTON, V. (1957). Perceptual and personality dynamics in reactions to ambiguity. *Brit. J. Psychol.* **48**, 200-15.

- HAMILTON, V. (1960). Imperception of phi: some further determinants. *Brit. J. Psychol.* **51**, 257-66.
- HOLZMAN, P. S. (1954). The relation of assimilation tendencies in visual, auditory, and kinesthetic time-error to cognitive attitudes of levelling and sharpening. *J. Personality*, **22**, 375-94.
- JONES, E. E. & BRUNER, J. S. (1954). Expectancy in apparent visual movement. *Brit. J. Psychol.* **45**, 157-65.
- KAISER, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.
- KAPLAN, J. M. (1952). Predicting memory behavior from cognitive attitudes toward instability. *Amer. Psychol.* **7**, 322. (Abstract.)
- KLEIN, G. S. & SCHLESINGER, H. J. (1951). Perceptual attitudes toward instability: I. Prediction of apparent movement experiences from Rorschach responses. *J. Personality*, **19**, 289-302.
- LOOMIS, H. K. & MOSKOWITZ, S. L. (1958). Cognitive style and stimulus ambiguity. *J. Personality*, **26**, 349-64.
- MARTIN, B. (1954). Intolerance of ambiguity in interpersonal and perceptual behavior. *J. Personality*, **22**, 494-503.
- SNYDER, R. (1956). Personality characteristics associated with the Kohnstamm phenomenon. Unpublished doctoral dissertation. University of Kansas.
- THURSTONE, L. L. (1947). *Multiple-factor Analysis*. Chicago: University of Chicago Press.
- TOCH, H. & ITTELSON, W. H. (1956). The role of past experience in apparent movement: A revaluation. *Brit. J. Psychol.* **47**, 195-207.

(Manuscript received 14 July 1960; revised 31 October 1961)





# THE DYNAMIC STRUCTURE OF ATTITUDES IN ADULTS: A DESCRIPTION OF SOME ESTABLISHED FACTORS AND OF THEIR MEASUREMENT BY THE MOTIVATIONAL ANALYSIS TEST

By R. B. CATTELL AND J. HORN

*Laboratory of Personality Assessment, University of Illinois*

AND H. J. BUTCHER

*Department of Education, University of Manchester*

In a systematic programme of research, two sets of factors have been repeatedly confirmed. Those of the first type are referred to as dynamic factors and comprise (a) drives or 'ergs' and (b) sentiments. These have been found when a wide range of adult attitudes has been measured by disguised tests and the results inter-correlated and factor analysed. The factors of the second type are referred to as motivation components and are derived from the measurement of one attitude by many different devices.

The number and nature of both types of factor are reviewed, and an experiment described in which a test battery was assembled so as to represent the principal factors of both types as economically and conveniently as possible. Thus for each expected dynamic factor two or more attitudes were included, all these attitudes being measured by four separate objective devices. The four objective devices were chosen to represent four of the most important components.

From the results of this experiment, the Motivational Analysis Test was constructed, designed to measure, in a short testing time, ten drives and sentiments, each by four devices. The test battery is clinically and experimentally oriented, to cover such drives and sentiments as fear, narcissistic sex, parental sentiment and the self-sentiment, for each of which a 'conscious, integrated' and an 'unconscious, unintegrated' score can be obtained, as well as a total score. Other established drives and sentiments with occupational and industrial relevance are included in another test battery still under construction.

Some practical uses of the Motivational Analysis Test are briefly suggested.

## I. INTRODUCTION

Systematic research over the last decade has resulted in considerable crystallization of findings about the major dynamic factors in adults, but wider and more effective use of these findings, for example in experimental study of the origins and consequences of drive tension levels, or in educational studies of the development of emotional integration and the self-sentiment, has only recently become possible now that a convenient measuring instrument has been constructed. The object of the present article is to provide an up-to-date survey of some scattered evidence and to describe the principles on which this new instrument has been constructed.

The research to be summarized differs in two important respects from most earlier work on attitudes, drives and sentiments. First, it differs from the approach of Drever (1917), Freud (1949), Garrett (1939), McDougall (1932), Murray (1938), and others, who have theorized about human drives on a basis of general and clinical observation, for it investigates structure by the multivariate analysis of quantitative manifestations. In this respect the difference is paralleled in work in animal psychology by the difference between the naturalistic approach of say, Lorenz (1952) and Tinbergen & Perlick (1950) on the one hand, and the quantitative, experimental approach

of, say, Anderson (1938), Beach & Jaynes (1954), Harlow (1955) and Haverland (1954) on the other. Secondly, the instruments used are disguised, objective devices, and not the questionnaires of traditional attitude measurement, which depend heavily on self-evaluation in the respondent.

A few years ago the present senior author surveyed the factor analysis of dynamic data available at that time, concentrating mainly on objective device studies, but also discussing outstanding questionnaire studies, such as those of Eysenck (1944), Guilford, Christensen *et al.* (1953) and Torr (1953). Table 1 summarizes some relevant researches, including unpublished work by Butcher and Sweney, and shows that sixteen dynamic factors have now been frequently replicated in separate studies, excluding the factors that have been found only in children.

Table 1. *Summary of dynamic structure replications to 1960*

The columns represent separate researches with authors as mentioned in the text, each X indicating a replication of the factor pattern.

Ergs (with associated emotion)	Cattell	Miller	Cross	Baggaloy	Sweney	Butcher
Escape (fear)	X	X	X	X	X	X
Mating (sex)	X	X	X	X	XX	X
Self-assertion (pride)	X	X	X	X	XX	X
Gregariousness (loneliness)	X	X	X	X	XX	X
Appeal (despair)	X	X	X	—	—	—
Exploration (curiosity)	X	X	—	X	XX	—
Protectiveness (pity)	X	X	—	X	XX	—
Narcissism (sensuousness)	X	—	—	X	XX	—
Play phantasy	—	—	X	—	XX	—
Rest-seeking (sleepiness, fatigue)	—	—	X	X	—	—
<b>Sentiments</b>						
Self	X	X	X	X	XX	X
Super-ego	—	X	—	—	XX	X
Religion	—	—	—	X	XX	X
Career	—	—	—	X	—	—
Sports and Games	—	—	—	X	—	—
Mechanical	—	—	—	X	X	—
<b>Suggestions of erg and sentiment patterns appearing in child studies only</b>						
<b>Ergs</b>						
Constructiveness (creativity)*	—	—	—	—	X	X
Pugnacity (anger)	—	—	—	—	X	X
Acquisitiveness	—	—	—	—	XX	X
Self display	—	—	—	—	XX	X
<b>Sentiments</b>						
Patriotism	—	—	—	—	XX	X
Amusements	—	—	—	—	XX	X
Aesthetic	—	—	—	—	X	X
Adult aspiration	—	—	—	—	XX	X

\* In this case we are dealing with what may be an alternative interpretation of the mechanical interest sentiment.

It will be recognized that both the existence and the form of several of these dynamic factors are now known at a good level of confidence, though their interpretation, and particularly the relative influences of heredity and environment, need further debate and experiment. It is interesting that in many cases where a factor corresponds



approximately to a Freudian or McDougallian concept, the slant of meaning is also slightly different. For example, the narcissistic sex factor has an unexpectedly general character of self-indulgence (wanting comfort, sleep, good food, tobacco, alcohol), while much of the variance in pugnacious, destructive attitudes is absorbed into the fear factor.

## II. THE CHOICE OF DEVICES TO MEASURE STRENGTH OF ATTITUDE AND MOTIVATION

Extensive research has been concentrated on the investigation of at least a hundred objective devices for measuring motivation strength. Such investigation stands to dynamic structure investigation as analysis by latitude does to longitude, for whereas the dynamic factors (drives and sentiments) appear among the correlations of attitudes, each measured once, the motivation factors emerge from the correlations of many devices, each measuring one and the same attitude. Since confusion may easily arise on this point, we consistently describe the drives and sentiments as *dynamic structure* factors and the factors derived from devices as *motivation component* factors.

Five to seven motivation components have consistently been found in both adult studies (Cattell, 1956; Cattell & Baggaley, 1956*b*) and among children (Cattell, Radcliffe & Sweney, 1962), and it is in terms of these that the validity of any device for measuring attitude strength should be evaluated. The task of having to validate in terms of seven such factors may seem disconcerting, but fortunately the seven oblique primary factors (labelled *alpha* to *eta*) resolve largely into two broad and some minor second-order components (Cattell, 1959). The two most important second-order factors have been called the Integrated, Conscious Motivation Component, *A*, and the Unintegrated, Unconscious Component, *B*, in accordance with the clear grouping of the primaries. For example, the three first-order components *alpha*, *delta* and *epsilon*, interpreted respectively as id, physiological responses and unconscious complexes, all load on the *B* second-order factor. Consequently, if one uses a battery of different devices so designed that some are known to load on *A* and some on *B*, any attitude can now in principle be given two scores representing *A*, the conscious, and *B*, the unconscious component, or these can be added to form a single total measure of attitude strength.

The experimental possibilities thus opened for investigating the theoretical concept of motivation strength are interesting, but our purpose in describing the motivation component structure at this point is to explain the basis on which devices were selected to form the most valid possible battery. The intention also was to produce a battery which would measure a given attitude in little more than the 5 min. traditionally required for a self-evaluation questionnaire of about twenty items.

Our selection from some 100 devices available (Cattell, 1957) was governed by the following aims: (1) to have a battery of up to six subtests per attitude, no more than two of which would need to be administered individually with apparatus; (2) to have the highest possible factorial validities in terms of both the integrated and unintegrated second-order factors, and at the same time to produce a meaningful total measure of attitude strength; (3) to represent also with tolerable distinctness even the primary components, *alpha*, *beta*, etc., in case an experimenter should wish to make as good an estimate of these as a single test could give; (4) to have tests with satisfactory reliability, or, failing this, tests that could be made satisfactorily reliable by

improved instructions or by increasing test length; (5) to have a test instrument that would prove psychometrically 'hardy' in test administration in the light of our several years of use of these devices; (6) to be able to use the battery in both R-technique (single administration) and P-technique (repeated administration); (7) to avoid very complicated scoring, and in particular to have as few devices as possible needing ipsative scoring (Cattell, 1957; Clemans, 1956).

To find tests simultaneously satisfying the above requirements demanded much search. For example, let us consider the first two motivation components, *alpha* and *beta*, which we wished to represent in the battery as adequately as possible by one device apiece. Among group tests the four of highest validity (mean loading over three studies) for the *alpha* motivation component were: *autism*, as exemplified in the misbelief device; *fluency*, or cues to interest action; *choice of means for ends*; and *recall of exposed material* (either on the short or long-term devices, and on cues or consequences). For the *beta* component the highest validities, among principles realizable in group tests, were: '*utilities*' (information on means to ends), *inhibition of distraction* (in the device of penetrating auditory distractions to conversations), *word association*, and *projection* (in the 'choice to explain' sentence completion device). Similarly loadings on the second-order components were carefully scrutinized and taken into account. Details of the tests referred to and of their loadings on the motivation components are provided by Cattell (1956, 1957). In addition, a compendium of objective tests of personality and motivation is in course of preparation by Cattell and Warburton, which will consolidate scattered references to a great variety of devices, and by describing systematically methods of administration, factor loadings and other particulars, provide, it is hoped, a comprehensive manual for research workers in this field.

Finally, among the group tests, *inhibition of distraction* was rejected, as requiring the removal of individual differences in sensory acuity; *recall*, on the grounds that it required too much skill in administration and scoring; *fluency*, on the grounds that the scoring was necessarily to some degree subjective; and *projection* as not being suitable for P-technique studies, since once the subject detects that his choice, not his explanation is wanted, the test changes its meaning.

On the individual tests, although *decision time* (on a paired comparison preference test), *psychogalvanic response*, *pulse* and *systolic blood pressure changes* give the best loadings, it is unreasonably taxing for subjects to have the arm cuff inflated a hundred times in the course of a 2 hr. experiment, and until medicine discovers a better way of measuring blood pressure than by stopping the circulation, this device can hardly be recommended for what is intended as a routinely applicable battery. It is also unfortunate that PGR and decision time data cannot be simultaneously gathered from the stimuli used while the group tests are given. This economy has proved possible in some circumstances, but generally only in conjunction with a questionnaire as distinct from an objective test device.

Thus, to summarize, four group tests were selected for the routine situation: *Information*, *Word Association*, *Utilities* and *Autism*, to give them their brief names, each representing a different motivation component. For research, this MAT battery has firstly an extension including three individual tests (*Decision Time* and the *Psychogalvanic* devices), and secondly a possible wider extension to twelve subtests



in all. The latter, to be used when maximum reliability and validity are required, is assembled by adding to the above seven subtests the devices labelled *Recall*, *Fluency*, *Systolic blood pressure change*, *Auditory distraction* and *Pulse change*.

The present research, however, operated with the core of four devices only, constituting the most widely usable MAT. Each device had six to nine items per attitude, and the four subscores were converted to standard scores and added, therefore, with equal nominal weights and without ipsatization. It would also be of interest to factor separately the same set of attitudes measured respectively on the unintegrated, or unconscious, and the integrated components, but our initial analysis is of one single interest score, nominally giving equal weight to integrated and unintegrated components, for each of the 28 attitudes.

Each of the attitudes to be measured (see Table 3 below) was first written out in the standard form (Cattell, 1956):

'I want to such an extent to do this with that',

to clarify the exact course of action around which the stimuli to be presented in the items must be designed to cluster. However, no attempt at item analysis within the attitude scales was made prior to the factor analysis, since our principle of validation required that such item analysis should be carried out against the dynamic structure factors found later in the analysis.

### III. PLAN OF THE EXPERIMENT

Research generally has an exploratory stage where the hypothesis is no more specific than that some structural relations exist to be recognized—and a stage of consolidation—where particular hypotheses generated from observed relations are to be checked. Ten years of systematic experiment by the senior author and his associates have brought dynamic structure research with objective measurement to a good degree of precision. The present research plan was, therefore, to check not on all the known factors, but only on a selected set likely to constitute a battery of greatest use for clinical and experimental research. In fact, the re-determination was necessary mainly because (a) the choice on utility grounds might not coincide with the list of most strongly replicated and confirmed factors; (b) the subsequent test construction, in any case, required fresh determination of the factors against which test items were to be validated in the sample.

To meet the demand for a relatively simple battery we decided not to exceed ten factors. (Many experimentalists may be interested only in *one* particular ergic tension and some clinicians seem interested only in the self-sentiment and super-ego, while all test users clamour for test brevity!) Since, as Table 1 shows, some 25 factors are known and comprehensive researches may need measures for all, the most realistic compromise seemed to be the construction of two overlapping batteries, each of ten factors and together covering 18 of greatest importance. This division was also designed to provide two instruments for rather different purposes, one being framed with a more clinical, and the other with a more occupational orientation to motivation assessment. The interest-adjustment test (IAT) which constitutes the latter will be described elsewhere. The immediate practical outcome of the present research is called the motivation analysis test (MAT) and is clinically and experimentally oriented.



A survey of clinical opinion suggested the inclusion of five ergic tension measures and five sentiment development measures, and taking also into account the current motivation measurement needs of learning theorists and students of perception, we finally settled on the list in Table 2.

Table 2. *Factors chosen for representation in the motivation analysis test*

Ergs	No. of attitudes representing each	Sentiments	No. of attitudes representing each
Sex	2 (nos. 3 and 4)*	Self sentiment	8 (nos. 21—28)
Self assertion	2 (nos. 5 and 6)	Super ego	4 (nos. 17—20)
Narcistic sex	2 (nos. 7 and 8)	Sweetheart-spouse	2 (nos. 15 and 16)
Fear	2 (nos. 1 and 2)	Parental home	2 (nos. 13 and 14)
Sadism-pugnacity	2 (nos. 9 and 10)	Career	2 (nos. 11 and 12)

\* These are the defining numbers of the attitudes in Table 3.

The chief problems in this list arise from the clinicians' request for three variables—sadism, interest investment in spouse, and attachment to parental home—which have not clearly emerged as unitary factors in earlier research. Although clinicians, e.g. Freud (Freud, 1949), confidently allege a sadistic need, attitudes representing destructive tendencies and pain infliction for its own sake (or as algolagnia) had failed to appear in the studies of Cattell & Miller (1952) and Cattell & Baggaley (1958), in the sense of clearly loading a single factor. None could be found either with the character of pugnacity (a term more cleanly separable from assertion than is the popular term 'aggression') or of sexual sadism. Pugnacious attitudes, directed to the destruction of frustrating objects, loaded principally the fear erg and possibly a sentiment recognizable as love of country. The theory had been tentatively stated that pugnacity is a unique ergic structure in having no independent energy of its own, by drawing its energy only from whatever other erg is frustrated (the need for security, therefore, by inference, being the most frustrated in our age). However, in these cases we agreed to try again, devoting particular care to the choice of attitudes (to represent both sadism-pugnacity and sex-sadism concepts) and of the items for these attitudes.

From the standpoint of good factor analytic technique, it is always necessary to represent each hypothesized factor by a minimum of two variables (attitudes here), and because of the special interest of the super-ego and self-sentiment the number was raised for these from two sentiments respectively to four and eight. The stimulus-response paradigm for an attitude has been introduced in the last section and it will be seen that the variables listed in Table 3 in every case conform to this concept. In most cases these attitudes are exact reproductions of those previously used with the same titles in earlier studies by the senior author and his associates.

It will be recalled that each of the above attitude scores is a total of scores on four devices—*Information*, *Autism*, *Word Association* and '*Utilities*' as described in the previous section. Together they represent for each attitude a total of 24 to 32 items and a testing time of 6 min. The total testing time was 2 hr. 40 min.

An important aspect of the present experiment was that the measures should be administered to a less homogeneous group than the young adult males employed in the earlier studies. For in the design of a generally applicable test it becomes necessary

Table 3. List of 28 attitudes hypothesized to mark 10 dynamic factor structures

Dynamic factor	Brief designation of attitude	Statement of complete paradigm
Escapo erg (fear)	1. Fear Atom-bomb	I want my home better protected against the terror of an atomic bomb attack
	2. Reduce death	I want to see the danger of death from disease and accidents reduced
Mating erg (sex)	3. Fall in love	I want to fall in love with a beautiful woman—a handsome man
Self-assertion erg	4. Satisfy sex need	I want to satisfy my sexual needs
	5. Smartly dressed	I want to be smartly dressed with an appearance that commands admiration
	6. Increase status	I want to increase my salary and social status
Narcissism erg (sensuous)	7. Easy time	I want to lie in bed in the mornings and have a very easy time in life.
	8. Enjoy delicacies	I want to enjoy fine foods, fine drinks, candies and delicacies
Sadism erg (sex?)	9. Destroy enemy	I want my country to go all out to destroy the enemy
	10. See violence	I want to see movies or plays showing gangster fights and violence, where many people are injured or slain
Career sentiment	11. Learn skills	I want to learn more about the technical skills in my job to be
Parental sentiment	12. Stick to job	I want to stick my job or chosen career
	13. Proud of parents	I am proud of my parents and want them to be proud of me
	14. Parents guidance	I want to turn to my parents for affection, comradeship and guidance
Sweetheart sentiment	15. Time with sweetheart	I want to spend time with my sweetheart, enjoying our common interests
	16. Gifts for sweetheart	I want to bring gifts to my sweetheart, to share in his or her delight in them
Super ego sentiment	17. Satisfy duties	I want to satisfy my sense of duty to my country and my God
	18. End vice	I want to see the end of gambling, idleness, excessive drinking, prostitution and all other forms of vice
	19. Be unselfish	I want to be unselfish in all my acts
Self sentiment	20. Avoid sinful sex	I want to avoid sinful, improper occasions of sexual expression
	21. Good reputation	I want to maintain a good reputation and a common respect in my community
	22. Normal sex	I want a normal, socially acceptable relation to a person of the opposite sex
	23. Look after family	I want to look after my family so that it reaches approved social standards
	24. Proficient career	I want to be proficient in my career
	25. Control impulses	I want to keep my impulses under sufficient and proper control
	26. Self respect	I want never to damage my sense of self respect
	27. Be sane	I want never to lose my mind (become insane)
	28. Know self	I want to know myself better



to demonstrate that the structure holds over the range of intended use, as well as to have items validated in the wider population. Accordingly, the sample used was one of 117 males ranging in age from 17 to 43 (28 in high school, 77 in college; 12 semi-skilled and skilled workers), 100 females, aged 16-58 (21 in high school; 45 in college; 34 semi-skilled and skilled workers). Thus the sample of 217 subjects was relatively heterogeneous in its sampling of the adult population, and at least gave sufficient range in sex, age and educational level to ensure that the factor structure found need no longer be considered restricted to young adult males.

Each subject did all the tests on a single day, since in dynamic, as distinct from ability and personality tests one deals with a picture which may change from week to week with changing stimuli and satisfactions.

#### IV. THE RESULTS AND THEIR ANALYSIS

Scores of the 217 subjects on the 28 attitudes were intercorrelated by the product-moment coefficient and factored by the centroid method. The number of factors was decided first by the method originally suggested by Ledyard Tucker and described by Cattell (1952, p. 298) and secondly by the method of using unit communalities and ceasing factoring when the first latent root fell below unity. The former method indicated 12, the latter 11 factors. The estimated communalities on the assumption of 11 factors converged to fixed values after 11 iterations of the factor analysis (fixed to the extent of not more than one point of change in the second decimal place.)

Since we had hypothesized ten factors, and obtained a statistical indication centring on 11, we rotated 11 on the assumption that if the eleventh were really unnecessary, it would collapse in the rotation. After ten visual rotations beyond an analytical 'oblimax' position, which raised the hyperplane count (i.e. percentage of zero, or negligible loadings in the whole matrix) from 50 to 61 %, a practically unimprovable simple structure position was obtained, since the hyperplane count (in the reference vector system) had remained level for the last three rotations. This can be considered a decidedly good simple structure in terms of standards of published studies and in terms of statistical significance by a modification of Bargmann's (1953) criterion. The significance values of the 11 factors are quoted in the last line of Table 4. Thus self-assertion, the self-sentiment and the career sentiment appear to have the less satisfactory hyperplanes, which might arise from these naturally having so central an influence in the whole personality that few variables are completely uninfluenced by them. In the case of the self-sentiment an insignificant hyperplane was unavoidable because we deliberately made eight out of 28 variables to be markers. The important thing here is that this is the *best* hyperplane, approximately orthogonal to the other factors, obtainable in a prolonged analysis after eleven rotations.

Accordingly, Table 4 presents the accepted test factor resolution. The attitude variables are in the order presented in Table 3, each designated briefly by a couple of words, and the factors have been re-ordered from the accidental sequence of their emergence into the order of their original statement, so that high values should run down the diagonal. The diagonal loadings which should, according to theory, be significant are in heavy type, as are the loadings outside these positions which exceed 0.3 and are considered substantial enough to require an attempt at psychological explanation.

Two main facts at once emerge from Table 4: first, that the main variance falls as hypothesized in ten factors, and secondly, that the positions of the main loadings very substantially support the hypotheses (based on the previous series of studies as well as specifically introduced here in the notion of a sadistic erg) as to the nature of the dynamic structure in the adult of our culture. After this, however, one observes also certain unhypothesized connexions—in the 'off-diagonal' loadings shown in bold—and also two instances of an hypothesized connexion failing to materialize. The latter are the failure of 'easy time' on narcissism and of 'maintain family' on



self-sentiment, although the latter may be due to most of our subjects (students) not having to maintain families. Most of the unexpected loadings, however, are consistent with traditional clinical views as to the direction of sublimation for particular ergs,

Table 4. The dynamic structures emerging from attitude correlations

	1	2	3	4	5	6	7	8	9	10	11
1. Fear atom bomb	46	-16	-24	-04	36	08	14	-10	-14	00	21
2. Reduce death	58	00	-13	10	-12	13	-07	02	08	-10	10
3. Fall in love	04	54	07	06	07	-20	05	-13	07	07	-03
4. Satisfy sex	-16	64	11	08	-12	07	-38	03	-09	-06	00
5. Smartly dressed	-12	00	41	16	-14	-19	20	-14	01	-15	03
6. Increase status	+16	12	28	09	-17	20	-20	-09	01	10	03
7. Easy time	-11	-01	34	10	23	-16	-28	13	-05	32	-29
8. Enjoy delicacies	52	-18	12	44	07	02	-04	08	00	10	-15
9. Destroy enemy	-05	-16	12	03	57	04	-06	05	11	-05	00
10. See violence	01	22	-18	42	21	00	-19	01	-04	-17	07
11. Learn skills	04	01	21	14	01	73	-02	05	04	08	-06
12. Stick to job	10	05	-10	-43	-08	45	-14	-02	26	-08	-06
13. Proud of parents	-11	09	24	-01	-02	04	78	07	01	24	06
14. Turn to parents	10	00	08	-17	-19	-05	82	-01	-04	-01	-08
15. Time with sweetheart	-04	-13	-10	10	-01	-06	12	76	-02	-01	-06
16. Gifts for sweetheart	04	13	10	-06	01	00	14	83	01	00	08
17. Satisfy duties	05	-18	13	-10	-11	18	-07	17	44	-06	30
18. End vice	-04	12	-39	-10	00	11	-01	11	33	19	-12
19. Be unselfish	-04	07	16	-01	08	-02	09	09	62	02	-10
20. Avoid sinful sex	04	14	-17	06	-11	07	16	-07	53	-02	12
21. Good reputation	-13	-21	02	22	-16	-02	00	-16	15	32	07
22. Normal sex	12	43	01	-04	15	20	14	17	-04	32	-07
23. Look after family	14	-13	-04	-01	00	-05	22	-02	00	06	-43
24. Proficient career	32	10	57	10	18	17	16	-03	-11	38	06
25. Control impulses	24	-07	-03	-09	10	-22	-05	06	23	30	43
26. Self respect	-07	-08	-04	10	-08	11	04	07	-04	58	-02
27. Be sane	04	16	03	06	08	04	10	-02	19	50	09
28. Know self	36	12	01	-03	-07	-14	07	05	08	33	-14
Significance of factor: $P <$	0.10	0.05	0.16	0.001	0.007	0.16	0.05	0.05	0.003	0.16	0.001

Table 5. Correlations among ergs and sentiments

	1	2	3	4	5	6	7	8	9	10	11
1. Fear	—	27	13	-05	-01	-26	-42	-09	-06	21	-23
2. Sex	—	—	-23	07	-08	-40	-33	14	-40	27	15
3. Assertion	—	—	—	14	-10	09	-09	-07	23	-07	-31
4. Narcissism	—	—	—	—	08	11	11	-18	-12	-22	30
5. Sadism	—	—	—	—	—	31	18	-26	-05	-20	01
6. Career	—	—	—	—	—	—	00	15	06	-20	00
7. Parental	—	—	—	—	—	—	—	11	25	-30	08
8. Sweetheart	—	—	—	—	—	—	—	—	11	16	-09
9. Super-ego	—	—	—	—	—	—	—	—	—	-06	-27
10. Self	—	—	—	—	—	—	—	—	—	—	02
11. Puritan self	—	—	—	—	—	—	—	—	—	—	—

as follows. Fear appears to express itself in drive to success in career, and in indulgence in eating. Sex desire, not unnaturally, intrudes into the self-sentiment ideal of achieving a good adult sex adjustment. Self-assertion enters into success in career more than does the self-sentiment which career was intended to represent. Narcissism is opposed to the onerous 'stick to the job' attitude. Sadism loads fear of the atom

bomb (tied into a reaction formation?). To have an easy time is unexpectedly a part of the self-sentiment—unexpected at least by McDougall, but making sense in terms of the concept we have been developing from experiment (Horn, 1961). Finally the well-defined sentiment of attachment to parental home and values has an appreciable unexpected negative loading.

The eleventh factor, though provisionally named a residual, is nevertheless psychologically consistent and interpretable. It combines self-control, rejecting indulgences, working hard and satisfying a sense of duty. Future research must ask whether this is a fragment properly belonging in the super ego, or whether there exists a second *type* of self sentiment which might be called a 'Puritan Ethic Self-Sentiment' incorporating more puritan or stoic values than the main self pattern of our time. An even more likely hypothesis is that this is a pattern produced by and peculiar to the distinctly older subsample within our sample.\*

#### V. CONSTRUCTION OF THE MOTIVATIONAL ANALYSIS TEST BATTERY

(i) *Item analysis.* The results in the last section were in the nature of a final check on factorial structure. In general, with the partial exception of the narcissistic erg (probably on account of weakness of items) the factors found were as hypothesized. The next aim was to construct the test battery in the light of these results and in particular to select the most efficient items.

Several possibilities were considered at this point. The items in each attitude scale could be validated either against the pooled score for that attitude, or against the factor on which the attitude loaded. The second alternative was chosen, in the expectation that a majority of experimenters using the instrument would find a factor score more useful and easily interpretable, and would be more interested in, say, a tension level on the fear erg, than in score on a single attitude. Secondly, the restriction was imposed that no attitude should contribute to the score on more than one factor. Although several tests, notably the Guilford-Zimmerman, allow the same item to contribute to more than one factor, the fact is that there is always more specific variance than common factor variance in an item, with the attendant danger of producing spurious correlations among the factor scores to a greater degree than is justified by the gain in factorial validity. Accordingly, the next step was to estimate for the 217 subjects factor scores on the ten factors the MAT is designed to measure, using Thomson's (1949) method, as discussed by Cattell & Baggaley (1956a) and programmed for an electronic computer by Hurley (1957).

Enough items had originally been included to allow for a considerable loss by attrition—about 26 items per attitude (six to nine in each of four devices). To assess the relative validities, it was necessary to employ the technique of parcelled factor analysis (Cattell, 1952). In this technique face-valid scales or parcels are first made up which are not too numerous for determining the factor structure. After factoring they are then 'undone' to permit the constituent items to be validated against the factors. In the present case, rather over half of the original items, though having in most cases

\* The unrotated matrix, transformation matrix, etc., on which the above is based have been deposited with the American Documentation Institute, whence microfilm may be obtained by reference to number 6572. The correlations among the oblique factors are presented here in Table 5, since they are important for theoretical developments on the relations of ergs and sentiments.



some positive relation to the particular criterion, were dropped from the several factor scales.

The items remaining generally correlated about  $+0.2$  or  $+0.3$  with the relevant criterion, which seems to be, in our experience, about as satisfactory a result as can be achieved, allowing for the heterogeneity of the test devices and the need to maintain suppressor action on unwanted factors.

(ii) *Reliability and factorial validity.* The final MAT as published (Cattell, Horn & Sweney, 1961) was still further modified by the inclusion of a few items in place of the weaker ones on two or three factors. The latest information on reliability and validity is given in detail in the *Handbook*, but further work is required with new samples. The internal consistency of the ten factor scales as determined by the Kuder-Richardson formula 20 ranges between 0.30 (sentiment to parental home) and 0.76 (self-sentiment), with a mean of 0.47. These figures seem low, for two reasons. First, the fact that each attitude is measured by four separate devices inevitably works against homogeneity. Secondly and relatedly, the Kuder-Richardson coefficient (being in effect the mean of all possible splits) is probably not the most appropriate measure. One should perhaps have computed a 'herring bone' type of coefficient by randomly splitting subjects' scores as measured by each of the four test devices, summing the four sub-scores and correlating with the other form, to obtain the most realistic estimate.

Factorial validities of the ten scales were in fact obtained by an analogous method and ranged between 0.36 (sentiment to parental home) and 0.85 (self-sentiment). When the test is lengthened by the construction of an alternative form, it is probable that these figures can be raised.

## VI. SUMMARY AND DISCUSSION

### (i) *The purpose and scope of the MAT*

The technique of factor analysis, generally with 'blind' rotation of the factors to oblique simple structure, has been adapted in a series of experiments over the last decade to the analysis of interests and attitudes, measured by relatively objective devices. Of the factors that have repeatedly emerged, some ten have been identified as sentiments or 'ergs'. (The term 'instinct' has been avoided, as having undesirable and quasi-philosophical overtones.) Another six to nine have been identified as sentiments, or acquired patterns of expression apparently corresponding to sociological institutions.

The research reported marks a change of emphasis in a systematic programme. The earlier research was theoretical, being devoted to establishing some basic patterns of human motivation and to sifting a large number of possible measuring devices. The present aim is to incorporate the principal factors in true dimensional scales of known reliability and validity. The Motivational Analysis Test represents the first fruits of this policy, and measures ten important concepts, another ten (with some overlap) being measured by a second research instrument under construction.

Twenty-eight attitudes are measured by four test devices. These were chosen in the light of several earlier researches as markers for five ergs (sex, fear, assertion, narcissism and sadism) and for five sentiments (those of self, career, super-ego, parental home and





# SKILL AND JUDGMENT OF FOOTBALLERS IN ATTEMPTING TO SCORE GOALS

## A STUDY OF PSYCHOLOGICAL PROBABILITY

By JOHN COHEN AND E. J. DEARNALEY

*Department of Psychology, University of Manchester*

The paper describes an experimental study of certain aspects of football in the context of psychological probability. Two First Division teams, a University team and a Grammar School team acted as subjects. The methods employed include extensions of previous techniques and new applications. A further analysis of certain experimental situations was made by an examination of films of actual matches. Supplementary evidence was obtained from the statistical records of First Division football clubs and from discussions with their managers. The results reveal systematic relationships between skill and judgment in goal scoring. Theoretical assumptions and implications are considered.

### I. INTRODUCTION

The investigation to be described is an attempt to apply concepts and methods we have devised for the general study of skill and judgment in risk-taking to characteristic situations in football, for the footballer's play, like athletic activity generally, involves risk in the sense that it is a form of behaviour in uncertainty (Cohen, Dearnaley & Hansel, 1956, 1958a).

The question may be asked whether it would be possible to study football in terms of game theory. This, in our view, would not be feasible because skill, in whatever form it is manifested, including skill in football, may be regarded as the exercise of a connoisseurship the ingredients of which cannot all be made explicit (Polanyi, 1958). The game of football is normally so fast, furious and complex an affair that it precludes a detailed and systematic specification of each player's countless actions and judgments *in actual play*, as seems to be required by game theory. The fact that football skill cannot entirely be made explicit is an obstacle to the use of game theory but not to the application of psychological probability. For the specifiability of *all* the elements in skill is not a necessary condition for the study of football in terms of psychological probability, although it is desirable experimentally to render as much as possible explicit and specifiable. We may add that the traditional study of skill has been almost wholly based on *performance*. We take the view that such studies can be enriched if the subject's experiences, rendered by him articulate, together with his judgments and decisions, are investigated *on a par with* his performance.

There is possibly another obstacle to the use of game theory in the study of football, namely, in respect to the question of the attractiveness of actions at different levels of uncertainty. Von Neumann & Morgenstern (1947) raise the question: 'May there not exist in an individual a (positive or negative) utility of the mere act of "taking a chance", of gambling, which the use of the mathematical expectation obliterates?' They answer this question by concluding that 'concepts like a "specific utility of gambling" cannot be formulated free of contradiction on this level'. If by a 'specific utility of gambling' they mean a preference for this or that level of

uncertainty *as such*, irrespective of the outcome, then the psychological assumptions built into their theory may not do justice to the facts. Thus, in football, as indeed in many other situations, people vary considerably in their preferences for acting at different levels of uncertainty. True, it is experimentally difficult to disentangle psychological probabilities from subjective utilities, but there seems no good reason to suppose that differences in preferences for uncertainties are entirely attributable to differences in subjective utilities.

The exclusion of the concept of a 'specific utility of gambling' seems plausible and legitimate to von Neumann & Morgenstern in terms of the 'system of psychology... now available for the purposes of economics'. But they nowhere specify what this 'system' is. Nor do they provide the specifications for the 'much more refined system of psychology' which would make it psychologically implausible and illegitimate to exclude the concept in question.

Theories of decision-making, risk-taking and conflict, as well as systems of learning theory and conditioning generally assume as their model of a man a 'rational' creature who is expected to behave as Buridan's ass. There are doubtless occasions when it is difficult to distinguish between human behaviour and the behaviour of Buridan's ass. But it is not necessary to assume that man is only capable of such behaviour. As elsewhere suggested (Cohen, 1960, 1961; Cohen & Cooper, 1961*a, b*), other models may prove more suitable, models which do not act in robot-like fashion but which resolve their conflicts by rendering their predicaments explicit and thus making it possible to engage in effective communication with themselves as well as with other 'players'.

The justification for an *experimental* study of football is based on the fact that the pattern of circumstances fluctuates from moment to moment, and appears different to each player. In order to establish principles or generalizations we must place the players in a simplified situation which remains constant and which is substantially the same for all players.

A simplified situation must necessarily be somewhat artificial and may therefore not evoke the players' full resources. Such an objection may be made against any experimental situation which is a simplified version of reality. We must weigh the advantages of obtaining relatively precise information in artificial circumstances against the disadvantages of unreliable and imprecise information in a 'real life' situation. There are, of course, *degrees* of experimental simplification and artificiality. The type of situation we have chosen for study is one which attempts to combine the advantages of the laboratory with those of a 'real life' situation, though inevitably sacrificing a little of each.

The two general questions we have set ourselves to examine may be put as follows:

- (i) What is the relationship between a player's *assessment* of his skill and his *actual* skill?
- (ii) What is the relationship between attempting to shoot a goal and (a) the player's distance from it; and (b) his confidence of success?

It should be clear that we are not attempting to measure levels of aspiration, i.e. what the player would like to achieve, but levels of *realistic* expectation (Cohen & Hansel, 1957). The relationship between them would be a third question which we are not here examining.



Our principal source of data for answering these questions is experimental. We have, however, supplementary sources of evidence from films of three matches and from statistical records.

## II. PROCEDURE

The experiments were conducted with players from two professional teams (Manchester United and West Bromwich Albion) in the First Division of the English Football League, the Manchester University first team, and the Manchester Grammar School first team. The investigations took place on their respective fields, which were of standard size. In three of the four teams, the regular goalkeeper was employed; in the fourth team the goalkeeper was a reserve. Each player was studied individually while playing against his goalkeeper, and no player was allowed to watch the experiment before he himself had acted as subject.

### *Stage 1: Assessments of skill*

The player stood well within his own half of the field, at a distance from the goal at which he thought he had no chance of scoring, and facing the goalkeeper who stood on his goal line. He was asked to imagine that the goalkeeper would come into play the moment he (the player) would begin to take action in attempting to shoot a goal. He was free to think of kicking a static or moving ball, whichever he preferred. Before making any attempt he was to assess his chances of success.

With the ball at his feet, each player was then asked to approach his opponents' goal down the centre of the field. He was instructed to stop when he *first* felt that he had a chance of scoring once in a hundred attempts. This point was marked by a small peg.

Accompanied by the experimenter, the player then continued to approach the goal while indicating the different points where he *first* felt he could score once, twice, three times, and four times in five attempts, and finally the point at which he *first* felt he could score ninety-nine times in a hundred attempts. These points were also marked by pegs. Before making these assessments he was told that he would subsequently make five shots at the goal from each of the six points he indicated. Thus, it was explained to him, the accuracy of his assessments would be measured.

### *Stage 2: Longest shot attempted*

Starting nearer his own goal than the point at which he thought he could first score once in a hundred attempts, the player, with the ball at his feet, ran towards the opponents' goal down the centre of the field. He was asked to imagine that he was playing in an important match, had 'broken through' and was being hotly pursued, but that only the goalkeeper stood between him and the goal. He was to shoot at the first appropriate moment. When he shot, the point at which he shot was marked by a peg. He then made four further and similar shots from this point. The success or failure of the shots was noted.

### *Stage 3: Measures of skill*

The player made five shots from each of the points at which, at Stage 1, he had estimated  $1/5$ ,  $2/5$ ,  $3/5$ ,  $4/5$ , and  $99/100$  successes, and their success or failure was noted.

Finally, the experimenter determined the furthest point from the goal at which the player succeeded five times in five attempts. This was necessary when any player had not scored 5/5 at any distance from the goal or when the distance between the points at which he felt he could score 4/5 and 5/5 was greater than 2 yards.

The experimenter always kept level with the player, so far as this was possible, and he was able, within a margin of error of one yard, to locate the spot from which each shot was made. The pegs were coloured black on one side and white on the other and were so placed as to be inconspicuous to the player as he approached the goal.

*Stage 4: Method preferred by player in order to feel most certain of scoring.*

In this situation, the goalkeeper stood on his goal line and the player at the centre of the field. The player was instructed to proceed in any manner he wished in an attempt to score a goal while trying to make as sure as possible that his strategy would be successful. The goalkeeper, on the other hand, was instructed to do his best to prevent a goal. Both came into action at a signal from the experimenter. The method of play chosen by the player was noted and the point at which he made his shot was marked. The final point to which the goalkeeper advanced was also noted. Two observers were necessary, one for the goalkeeper and one for the player.

After this stage was completed, the distances from the goal line of all the points marked throughout the experiment were measured (to the nearest foot).

The experiments took place in March and April 1960 and the number of players studied was 33, comprising 12, 8, 5 and 8 from the four teams respectively.

#### (i) *Assessments of skill*

### III. RESULTS

Before considering our first question (see subsection (ii)), we must examine the pattern of assessments. In Stage 1, the player indicated the distances from the opponents' goal at which he expected to achieve different proportions of success, once in a hundred attempts, once in five attempts and so on. In approaching the opponents' goal he was asked to stop at the successive points from which he thought he could *first* score in a given proportion of attempts. The distances from these points to the goal we take to be measures of the player's assessment of his skill in the given circumstances. It should be borne in mind, when comparing teams, that each team had its own goalkeeper. The fact that any chosen distance is not independent of distances previously chosen by a player should also be noted. The average distances for each of our four teams are shown in Table 1.

Table 1. *Average distances to the goal (in feet) corresponding to variations in expected success*

Team	Expected success						N
	1/100	1/5	2/5	3/5	4/5	99/100	
Manchester United	106	67	50	39	29	19	12
West Bromwich Albion	96	68	55	46	36	28	8
University	96	59	45	36	28	17	5
Grammar	116	71	55	45	33	19	8
Mean	104	67	52	42	31	21	—

On *a priori* grounds it would not be entirely unreasonable to assume that, in general, estimates of success would increase as the distance to the goal decreases and with the corresponding increase in angle subtended by the goal. The data in Table 1, the last row of which is plotted in Fig. 1, show the form of this relationship, which enables us to state the mean distances corresponding to particular expectations of success. The general form of the relationship characterizes each individual player, and our data make it possible to specify for each player the particular distance corresponding to his varying expectations of success. All the players stated that any point nearer the goal than the one corresponding to an estimated 99 per cent success would diminish their chances of success because the goalkeeper would then be in a more advantageous position.

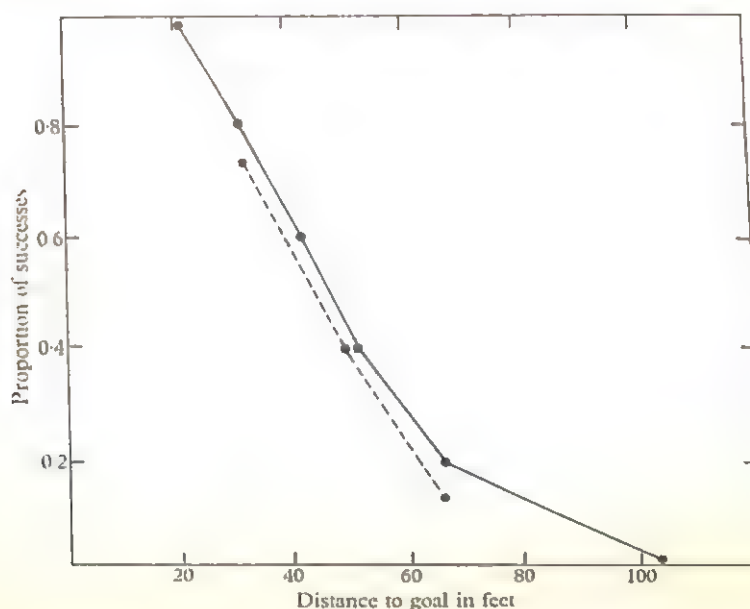


Fig. 1. Estimated and actual success at different distances to the goal.  
—, Estimated; ---, performance.

The variation between the rows in Table 1 is significant ( $P < 0.01$ ). This variation principally consists in the fact that, except when the distance from the goal is great the West Bromwich team reaches any given level of confidence at a further distance from the goal than any of the other three teams; this is particularly striking at the 99/100 level. Furthermore, in respect to a given degree of confidence being held farther away from the goal, the teams are in the following order: West Bromwich, Grammar School, Manchester United, University.

Thus, although there are significant differences between the teams, the differences are not characteristic of professional versus amateur teams. The University players are the most cautious in the sense that they have to be nearer the goal than players of other teams before they reach any particular level of confidence.

The standard deviations about the means shown in Table 1 are presented in Table 2.

An analysis of variance of Table 2 shows that the rows differ significantly ( $P < 0.01$ ). Two points emerge. First, the two professional teams combined vary



more among themselves than the two amateur teams combined. At the 1/5, 2/5 and 3/5 levels,  $P < 0.05$ . Since the means of the professional and amateur teams do not differ significantly, this suggests a greater heterogeneity of play among the professional teams. One reason for this difference might be the absence of players with powerful kicks among the amateur teams. There is one exception to this, the large variance of the Grammar School team at the 1/100 level which reflects its large mean (116 ft.), itself the result of very unrealistic estimates on the part of two schoolboys. Secondly, in general, standard deviations decrease as the distance to the goal diminishes, as is to be expected ( $P < 0.01$ ). If so, the magnitude of the standard deviations might be expected to reflect the magnitude of the means. This is so except in the case of Manchester United, which has larger standard deviations at the 1/5, 2/5, 3/5 and 4/5 levels than would be predicted on this ground.

Table 2. *Standard deviations (in feet) of the distributions of the distances, the means of which are shown in Table 1*

Team	Expected success					
	1/100	1/5	2/5	3/5	4/5	99/100
Manchester United	24	14	12	10	9	4
West Bromwich Albion	16	12	9	7	5	5
University	7	4	3	4	6	4
Grammar	20	5	6	5	6	5

The differences between the standard deviations of the teams might be interpreted to mean that the play of Manchester United fluctuates more widely than that of West Bromwich. This question of team variability is, however, more complex than appears on the surface. It would perhaps not be entirely safe to argue from the differences between players within a given team to differences between this team and other teams over a period of time. On the other hand, interchange of players in a team of greater variability might well produce more effect than a corresponding interchange in a team of less variability.

We have compared this variability of our two professional teams with the week by week changes in their place in the League table during the 1959-60 season. It is these changes which catch the public eye. The total number of such changes in position was summed for each First Division team over the season. The four teams with the fewest changes in this respect consist of the three teams which led at the end of the season and of the team at the bottom of the table. Manchester United had a larger number of changes than West Bromwich.

The final rank order of our two professional teams when it is based on week by week changes differs from the final order when it is based on the variance through the season of the number of points scored each week. On this latter scale West Bromwich fluctuates more, although the ratio of the variances does not reach significance at the 5% level.

This apparent inconsistency between the measure based on place changes and that based on points might be due to the fact that the variation between the teams is due to several independent factors and not attributable to a single source. Thus other measures might be based on an examination of (a) victory or defeat on the two

occasions that each team plays every other team; (b) sequences of victory (or defeat) as compared with successive alternations; (c) systematic differences between the first and second halves of the season; (d) changes in the composition of the teams; and (e) number of goals scored. An analysis of variance of such information, which is beyond our present scope, might, in fact, reveal such multiple sources of variation.

## (ii) Accuracy of assessments of skill

Each player made five attempts to score a goal at each distance marked at Stage 1. He was advised to perform to the best of his ability, and not to try and make his assessments accurate retrospectively. The degree of accuracy achieved is shown in Table 3. We have omitted to show the means at the 1/100 and 99/100 levels because at the former level only one player scored at all and at the latter level only five players failed to score at every attempt. It should be noted that at the 99/100 level the player can score less than 5/5 but he cannot score more than 5/5. Hence the mean performance is bound to fall short of 5/5. A similar argument, but in the other direction, applies to the 1/100 level. The means would therefore be misleading.

Table 3. Average number of successful shots in relation to expected success

Team	Expected success				N
	1/5	2/5	3/5	4/5	
Manchester United	0.3	1.5	2.5	4.2	12
West Bromwich Albion	1.1	2.1	3.1	3.4	8
University	1.0	2.4	3.0	3.6	5
Grammar	0.4	1.1	2.3	3.6	8
Mean	0.7	1.7	2.7	3.8	—

The assessments of success were only made while the player was *approaching* the goal because this procedure resembles actual play. Theoretically, the estimates could equally have been made when the distance from the goal progressively increased. The method adopted, as we may presume from previous experiments (Cohen *et al.* 1958*b*), has the effect of giving rise to more confident estimates at any given distance than if they had been obtained in the reverse or random order.

The means, expressed as fractions of unity, are plotted in Fig. 1. As may be seen, there is a close correspondence between estimates and performance: one third of the estimates are accurate. The expected performance is slightly overestimated, since 58 of the 132 estimates are overestimates, 29 are underestimates and 45 are accurate.

The differences between estimates and performances are much the same in the four teams, although there are signs that Manchester United overestimates its performance and that West Bromwich (to a less extent) underestimates it. It is perhaps no accident that the youngest team shows a significant mean error in assessment of skill in two out of four instances; in both cases, the error is an over-estimation of skill.

We have already given a reason for omitting from Table 3 the means at the 99/100 level. A measure of the accuracy of estimates at this level may be obtained by com-

paring the furthest distance at which the player thought he could nearly always succeed with the furthest distance at which he did, in fact, always succeed. This comparison is shown in Table 4.

Table 4. *Mean farthest distances (feet) at which players succeed 5/5 and believe they will succeed 99/100*

Team	Succeeds 5/5	Believes he will succeed 99/100	Difference
Manchester United	26	19	7
West Bromwich Albion	30	28	2
University	24	17	7
Grammar	25	19	6
Mean	26	21	5

It might be objected that such a comparison is not strictly accurate because the distance at which the estimate is made relates to 99/100 and the other distance to five successful shots in five attempts. We may reply that it was impracticable to ask the players to make a hundred shots and the players were reluctant to specify a point for five successful shots in five attempts.

At the 99/100 distance all four groups of players appear to underestimate their ability. Even the Grammar School players, on the average, think that in order to succeed 99/100 they must be, at most, 19 ft. from the goal whereas they succeed 5 times out of 5 when they are 25 ft. away. Much the same applies to the other teams. The players' assessments, however, may actually be more accurate than would appear from Table 4 since it may be easier to score 5/5 than 99/100.

The suggestion of overestimating at the further distances (Table 3) and of underestimating at the nearer distances (Table 4) is consistent with the results we have obtained in a number of other experiments (Cohen & Hansel, 1956*a, b*), namely, that people tend to overestimate their skill when the task seems to them hard and to underestimate it when it seems to them easy.

(iii) *Longest shot, i.e. farthest distance from the opponents' goal at which a player first attempts to score*

In previous investigations (Cohen *et al.* 1955, 1956, 1958*a*; Dearnaley, 1958) we have found a useful measure in the 'maximum risk-taking level', by which we mean the most uncertain task which a person is prepared to undertake. In the present context, this is represented by the level of uncertainty at the farthest distance from the goal at which the player will shoot (i.e. longest shot).

Table 5 presents the mean distances and standard deviations of the longest shot of the four teams and the mean psychological probabilities ( $\psi$ ). The latter are based on measures obtained for each player individually by interpolating in the relationship between his assessments and distance. The final column shows the corresponding performance of the four teams expressed as probabilities.

The variation between the mean longest shots is statistically significant ( $P < 0.01$ ); a difference of 15 ft. between the means is necessary for significance at the 0.05 level. The West Bromwich players differ from the other three teams in that they are



prepared to make their longest shot, on the average, farther away from the opponents' goal and at a lower level of confidence or, in other words, at greater uncertainty.

The mean maximum risk-taking level ranges from 0.08 to 0.33, which corresponds to statistical probabilities ranging from 0.10 to 0.16. The distances for the longest shot of each player are, with two exceptions, shorter than the 1/100 distances (see Table 1).

Table 5. *Mean distances at the longest shot, S.D.'s maximum risk-taking levels ( $\psi$ ), and corresponding performance*

Team	Longest shot	$\sigma$	$\psi$	Performance
Manchester United	59	11	0.29	0.16
West Bromwich Albion	83	14	0.08	0.10
University	51	14	0.33	0.16
Grammar	68	16	0.27	0.16

Table 6. *Margin of hazard and margin of safety*

Team	Margin of hazard (ft.)	Margin of safety (ft.)
Manchester United	33	35
West Bromwich Albion	53	56
University	27	39
Grammar	43	49

Further light on the characteristic feature of West Bromwich to which we have just referred is shed by two further comparisons, namely, (i) between the longest shot attempted and the longest shot at which the player succeeds 5/5; and (ii) between the longest shot attempted and the longest shot at which the player *thinks* he will succeed 99/100. The two resulting differences we call margin of hazard and margin of safety respectively; they are presented in Table 6.

The 'margin of hazard' indicates the greatest danger of failure which a person incurs. Thus, the order of the teams in respect to the greatest danger of failure they are prepared to incur is West Bromwich, Grammar School, Manchester United, University. The variation between teams in the margin of hazard is statistically significant ( $P < 0.05$ ); a difference of 16 ft. is necessary for this level of significance to be achieved.

The variation in margin of safety, however, just falls short of significance at this level. The margin of safety need not closely resemble the margin of hazard. They happen to be closely alike in the present instance because what the players *can* actually do is very similar to what they *think* they can do.

The general view among managers, so far as the longest shot is concerned, is that the better player would shoot 'late' if possible, despite the fact that he may be tackled within the penalty area (54 ft.). He would shoot at 75 ft. if the goalkeeper could be caught unawares, but never from more than 120 or 150 ft. In short, other things being equal, the better player would shoot nearer to the target. How near he would feel it necessary to get would depend on how he assessed his strength and accuracy of shot. Some managers expressed the view that the better player would not shoot at a distance less than 54 or even 36 ft.

(iv) *Players' preferred method*

At Stage 4, when the player was given an opportunity of demonstrating his preferred method of shooting a goal, the goalkeeper moved forward to meet the attack. We found that the players adopted one of three methods: (a) shooting at the goal when the distance to the goalkeeper was still considerable; (b) waiting until the goalkeeper was within a few yards and then making a hard shot along the ground; and (c) attempting to take the ball past the goalkeeper and then shooting at an undefended goal. The three methods are not, however, to be regarded as mutually exclusive in an absolute sense because a player who prefers the first method might find himself compelled to try the second or third methods. The methods adopted by the different groups of players are shown in Table 7.

Table 7. *Methods adopted by players*

Team	(a) Shoot from a distance	(b) Shoot close to the goalkeeper	(c) Attempt to pass goalkeeper
Manchester United	10	1	1
West Bromich Albion	4	1	3
University	1	0	4
Grammar	5	1	2
Total	20	3	10

There were 29 goals in 33 shots. It is clear therefore that when the player can take the initiative in this way he is in a decidedly advantageous position. The four players who failed to score comprised two from Manchester United and two Grammar School boys. One player from each of these teams failed in using the first method and one from each team failed in using the second. The majority of the managers preferred either the first or, to a slightly less extent, the third method.

(v) *Reliability of measures*

We have obtained an indication of the consistency of the measures from one professional player, internationally famous, who repeated the experiment a week later. His distances corresponding to his varying expectations of success on the two occasions are shown in Table 8.

Although the absolute differences (a) - (b) remain fairly constant at the varying levels of expected success, the percentage differences (final column) increase with increase in expected success. This may be attributed to the greater variability inherent in the latter situation, because the goalkeeper's actions may affect the player more as he approaches the goal. Perhaps the most important difference, from the point of view of reliability, is the one at 1/100.

On the two occasions, the five attempts at each distance never varied by more than one success. Our player's two longest shots varied by less than 2 ft.: they were 74 and 75 ft., respectively, a difference which lies within the errors of measurement. He chose the first method of approach on both occasions, the distances at which he shot being 40 and 45 ft., respectively.

Table 8. Distances (in feet) corresponding to variations in expected success

Expected success	First experiment (a)	Second experiment (b)	Difference (a) - (b)	Percentage difference $\frac{(a)-(b)}{(a)} \times 100$
1/100	114	121	-7	-6
1/5	83	83	0	0
2/5	70	61	9	13
3/5	54	48	6	11
4/5	40	35	5	13
99/100	32	24	8	25

## IV. SUPPLEMENTARY EVIDENCE FROM FILMS

(i) We have tried to determine, from films taken of three football matches, the number of attempts made at shooting a goal and the distance from the goal at which they are normally made. The three matches studied were the Football Association Cup Final between Manchester City and Birmingham City in 1956, the European Cup Final between Eintracht and Real Madrid in 1960, and a First Division League match between Manchester United and another team.

(ii) This last film was obtained by filming our team's play in the opponents' half of the field. The film was taken from a position above the field in the first row of the Grandstand, so that the white lines appeared in each frame of the cinefilm. The play was analysed by finding the frame in which the attacking player made an attempt at shooting a goal. His location on the field of play was determined by his relationship to the white lines. The accuracy of the estimated positions is subject to an error of approximately  $\pm 3$  ft.

In this First Division match our team made 22 attempts to shoot a goal of which 3 were successful. These 3 successful attempts and 13 unsuccessful ones were analysed in respect to distance and angle from the centre of the goal. The three successful attempts were made at small angles of  $35^\circ$ ,  $11^\circ$  and  $22^\circ$  respectively. As our experiments were made at an angle of  $90^\circ$ , it is not entirely proper to compare our longest shot (59 ft.  $\pm 11$ ) with the distances at which the three goals were scored i.e. 90, 33 and 60 ft. respectively, but two of these latter distances are less than our mean longest shot plus one standard deviation.

A situation arose during this game which closely resembled one of our experimental situations. A Manchester United player had made a break-through. The distance from the goal at which he shot (unsuccessfully) was 63 ft., which is approximately the distance in the experiment (Stage 1) at which he thought he could succeed twice in five attempts, although in the experiment he never succeeded at this distance. His longest shot at Stage 2 was 74-75 ft., and at Stage 4, his preferred method was to shoot from 40-45 ft. It happens that this player is the one for whom we have two sets of data set out in subsection (v) of the Results.

A detailed comparison between this film and our experiment is not possible in relation to individual performance because of the invariable presence of other players near any attacker. We may, however, compare features common to a number of players with the experimental results. Thus all attempts to shoot a goal



which were more or less perpendicular with the goal were within the penalty area, and at a distance corresponding experimentally to at least two successes in five attempts.

We recognize that most goals are scored from within the penalty area, and this is borne out by our own film data. Many goals are therefore scored not by straight shots but by 'side foot' efforts. In our experiment the players were perfectly free to kick directly at the goal or to proceed by 'side foot' efforts.

The actual positions on the field from which shots, successful and unsuccessful were made are shown in Fig. 2.

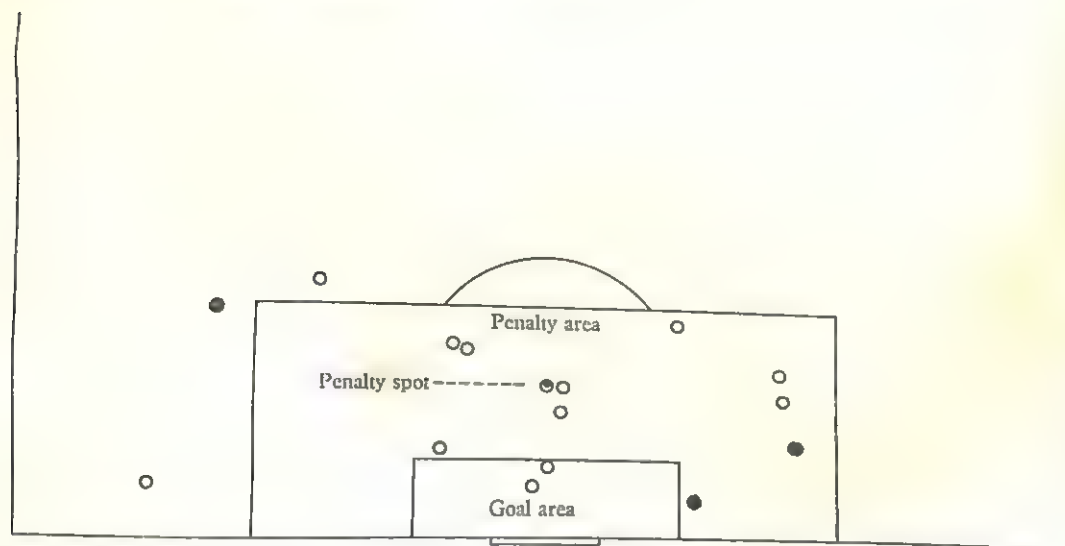


Fig. 2. Position from which shots at the goal were made by the first professional team in a League match. ●, Successful shot; ○, failure.

We remind the reader that the width of the goal is 24 ft., the goal area extends to 18 ft., the penalty spot is 36 ft. from the goal, and the penalty area 54 ft. from the goal. In the experiment the distance 18 ft. corresponds to an estimated success of approximately 99/100, the distance 36 ft. to 3/5 or 4/5, and 54 ft. to 2/5.

(iii) Unfortunately we were not able to analyse the two other films in the same way because they were not taken with our aims in mind. Nevertheless, we are able to draw certain conclusions with regard to (a) successful and unsuccessful attempts inside and outside the penalty area; (b) the number of attempts at shooting a goal in the two halves of the game; (c) attempts by winning and losing teams. The relevant information is presented in Tables 9-12.

First, we note the far greater frequency of successful shots within the penalty area than outside it. There were 13 successes out of 44 attempts *within* the penalty area and only one success out of 32 attempts *outside* it. In the European Cup Final, 7 of the attempts were made within the *goal* area, and of those 6 were successful. Thus, in actual play, as indeed in our experiment, the chance of scoring a goal increases as the attempt is made nearer the goal, as may be expected.

Secondly, the total number of attempts declined during the second half of the game in both matches, as shown in Table 10.

Thirdly, the winning teams make more attempts within the penalty area than the losing teams. Outside the penalty area, the number of attempts made by winning and losing teams is much the same (see Table 11).

Table 9. *Successful and unsuccessful attempts within and outside the penalty area in two Cup Final matches*

	F.A. Cup Final	European Cup Final	Total
Within penalty area			
Successful	4	9	13
Unsuccessful	15	16	31
Outside penalty area			
Successful	0	1	1
Unsuccessful	11	20	31
Total			
Successful	4	10	14
Unsuccessful	26	36	62

Table 10. *The number of attempts in the first half and the second half*

	F.A. Cup Final	European Cup Final	Total
First half	18	24	42
Second half	12	22	34
Total	30	46	76

Table 11. *Attempts within and outside the penalty area by winning and losing teams*

	Within penalty area	Outside penalty area	Total
F.A. Cup Final			
Winners	12	4	16
Losers	7	7	14
European Cup Final			
Winners	16	12	28
Losers	9	9	18
Total			
Winners	28	16	44
Losers	16	16	32

A further question may be asked: do the winners differ from the losers only in making more successful attempts? or in also making more unsuccessful attempts? We have already seen in Table 11 that the winners did make more attempts than losers. We can now see from Table 12 that in the European Cup Final the winners also made more unsuccessful attempts, and in the F.A. Cup Final, winners and losers make the same number of unsuccessful attempts.

We should add that, from our own experience in filming and analysing a match, it is on occasion difficult to decide whether a shot is an attempt at a goal or not, but this difficulty rarely occurs.

We cannot, of course, generalize from two matches. More evidence on these matters is required. There is, however, nothing in the evidence from the films that is in conflict with the results of our experiment.

Table 12. *Numbers of successful and unsuccessful attempts made by winners and losers*

Attempts	F.A. Cup Final		European Cup Final	
	Winners	Losers	Winners	Losers
Successful	3	1	7	3
Unsuccessful	13	13	21	15
Total	16	14	28	18

## V. GENERAL DISCUSSION AND CONCLUSIONS

In subsection (i) of the Results we have considered the relationship between expectation of success and distance from the goal. Since the width of the goal is fixed, the angle subtended by the goal must vary with distance. Expectation of success varies directly with the angle, though not necessarily in simple linear fashion, and inversely with distance.

Let us now assume that the width of the goal could be varied. Distance from the goal could be held constant while varying the angle, and vice versa, the angle could be held constant while varying the distance. Under such conditions it would be possible to specify the precise relationship (*a*) between distance and expectation of success and (*b*) between angle and expectation of success. The study of these relationships are of interest for the experimental psychology of visual perception and skill and we are undertaking further experiments to elucidate them.

In the game of football itself the width of goal is, in fact, fixed, and perceptual constancies may decisively affect play. Other variables too, may influence the player's judgment and performance. Thus there is the anticipation effect, to which we have already referred (see subsection (ii) of Results), which may apply to his action as much as to his assessment of success. And there may be factors which do not derive from the game as such though they may disturb the player, and the onlooker may attribute the disturbance to defective play. We have in mind the report of a manager about one of his star players who was involved in an *affaire d'amour* which disrupted his play.

Views among our managers appear to differ sharply on the question whether players perform better or worse in the actual game than during training. Possibly those who hold opposing views have generalized from biased samples and both views are, in fact, consistent with the Yerkes-Dodson law. Thus, on the one hand, there is the view that 'Some people, but very few, play better under tension, most play far worse'; and on the other, the view 'When there is more at stake the player will use his ability to the fullest'.

Another divergence between managers appears in the view that confidence or luck affects success. 'Confidence is everything', declares one manager; 'We have had more than our share of luck this season', asserts another, with equal assurance.



In general, the managers on the whole agree that the good player would be better able to assess his abilities correctly. They also think that he would tend to be more confident, and to require a higher degree of certainty of success before shooting than would a player not so good.

Our conclusions may be set out as follows:

(i) We have established, in the circumstances of our experiment, a precise relationship between (a) the psychological probability of scoring a goal and (b) distance from the goal.

(ii) In respect to the foregoing relationships, there are characteristic differences between the teams studied, but no significant contrast appears between the professional teams and the amateur teams. The University players appear to be the most 'cautious' of the four groups.

(iii) Although professionals and amateurs do not differ significantly in respect to the mean distances corresponding to particular levels of expected success (or psychological probabilities), they do differ in respect to heterogeneity. There is greater variation in the professional than in the amateur teams.

(iv) There is a close correspondence between estimates and performance in all teams, though there is perhaps a slight tendency to overestimate expected performance.

(v) It would seem that certain teams tend to overestimate their performance and other teams to underestimate it. If so, this might be due to the policies of the teams and is a matter that would repay further investigation.

(vi) The longest shot, i.e. the farthest distance from the goal at which players first attempted to score, ranged from a mean of 51 ft. in the University team to a mean of 83 ft. in West Bromwich Albion. The differences between the teams are significant.

(vii) The corresponding psychological probabilities (i.e. 'maximum risk-taking levels') ranged from a mean of 0.08 in West Bromwich to a mean of 0.33 in the University team. Thus the University players were not prepared to shoot unless they felt they had at least one chance in three of success, whereas West Bromwich were content with a subjective chance of eight in a hundred. In actual performance at their longest shot the first, third and fourth teams were alike in scoring at the rate of one in six attempts, and the second team at the rate of one in ten attempts. Thus, taking conclusion (vi) into account, the West Bromwich players differed from the other teams in being prepared to make their longest shot, on the average, farther away from the goal and at a lower level of confidence.

(viii) When the players were free to choose their preferred method (as in Stage 4), the methods chosen fell into three categories. Under these conditions the players were in a very advantageous position *vis à vis* the goalkeeper.

(ix) An indication of the reliability of the measures was obtained by a more detailed study of one player on two occasions.

(x) In the film taken by us we found that all attempts to score which were roughly perpendicular to the goal were within the penalty area and at a distance corresponding experimentally to at least 2 successes out of 5 attempts.

(xi) In the two Cup Final matches, there were 13 successes in 44 attempts *inside* the penalty area and one success out of 32 attempts *outside* it.

(xii) In these two matches, 42 attempts to score were made in the first half of the game and 34 in the second. The winners made more attempts within the penalty area, and they also made more unsuccessful as well as more successful attempts. Whether these differences are regular features of football requires substantiation from further observations.

#### APPENDIX

##### *Analysis of goals scored in relation to position in the team*

We asked the twenty-two managers to provide details of the number of goals scored in each position of the team for a period of 10 years, or since the team had been promoted to the First Division, whichever was the shorter period. Records were kindly made available by fifteen managers. Unfortunately all the information we sought was not available in four of the fifteen sets of records. We were thus left with records of goals scored in 3,192 First Division matches by 11 teams during 4 to 10 seasons, according to the team.

The total number of goals scored by the 11 teams in these matches was 5,348 (excluding penalties and goals by errors of opponents), representing an average of 1.7 goals scored by each team in each match. This information, in relation to the player's position, is set out in Table 13.

Table 13. *Goals scored from different team positions in 3,192 First Division matches during the decade 1949-59*

Position	Goals*	Per cent
Right back	16	0.3
Left back	22	0.4
Right half	96	1.8
Centre half	34	0.6
Left half	90	1.7
Outside right	566	10.6
Inside right	1086	20.3
Centre forward	1713	32.1
Inside left	1044	19.5
Outside left	681	12.7
Total	5348	100.0

\* Excluding goals scored from penalties or opponents' errors.

A number of features emerge from this Table.

- (i) Centre forwards scored the largest proportion of goals, almost a third.
- (ii) Next come the two inside forwards who scored about a fifth of the goals each.
- (iii) The outside forwards accounted for another quarter of the goals, the outside left scoring a statistically significant larger proportion than the outside right.
- (iv) The wing halves scored some 3% between them, the centre half scoring significantly less than either of these two. The latter's goal-scoring performance is on a par with that of the full backs, who contribute less than 1% of the goals.

The figures in Table 13 refer to the goals scored by eleven teams, but there is some variation from team to team. For example, the goals scored by the centre forward range from 24 to 39%, those scored by the inside left range from 11 to 25%, and those by the inside right from 10 to 40%. Three of these extreme values, 24, 11 and 40%, occurred in one team. The scoring potential in each position depends on the



policy (or 'strategy') of the team in relation to each particular player. A deep-playing centre forward may tend to score less but might compensate by making it possible for other forwards to score more. We should also bear in mind that an outstanding player in any position may alter the general pattern and symmetry of the game.

The figures for the forwards reflect the usual method of play. One team, namely that with the extreme values just referred to, was also peculiar in that the highest proportion of its goals was not scored by the centre forward.

The apparent goal-scoring superiority of the outside left, as compared with that of the outside right, requires some explanation. If most players are right-footed, they would presumably kick the ball more often to the left than to the right, and the ball would then make its appearance on the left of the field more often than on the right. On this assumption, the outside left would be provided with rather more opportunities than his colleague on the right. This hypothesis would, however, be untenable if in fact managers selected right-footed players to play on the right, left-footed to play on the left, and ambidextrous players for the positions of centre forward and centre half. The question at issue might be left for the zealous student of the game to investigate.

We are grateful to Mr Matt Busby, C.B.E., for his special advice and encouragement, to Mr Alan Hardaker and to all the managers of teams for their helpful collaboration and, in particular, to the managers and players of Manchester United and West Bromwich Albion. We are also indebted to the Manchester University team and to Lord James of Rusholme and his pupils at the Manchester Grammar School.

We also express our thanks to Mr L. MacDowall, manager of the Manchester City Football Club, and to Mr Peter Dimmock, C.B.E., of the B.B.C. Sportsview Unit for their courtesy in allowing us to examine their films. We are also indebted to Dr R. R. Skemp for his expert help in filming the First Division League match, and to our Technician, Mr J. P. Anderson.

#### REFERENCES

- COHEN, JOHN (1960). *Chance, Skill and Luck*. Harmondsworth: Pelican Books.
- COHEN, JOHN (1961). Man, Time and chance. *Proceedings of the Royaumont Colloque sur les Problèmes de la Personne*. (In the Press).
- COHEN, JOHN & COOPER, P. (1961a). Patterns of Preference in Equiprobable Situations. *Nature*, **190**, 231-2.
- COHEN, JOHN & COOPER, P. (1961b). A Model for Choice in Equiprobable Situations. *Acta Psychol.* **18**, 181-200.
- COHEN, JOHN, DEARNALEY, E. J. & HANSEL, C. E. M. (1956). Risk and hazard. *Oper. Res. Quart.* **7**, 67-82.
- COHEN, JOHN, DEARNALEY, E. J. & HANSEL, C. E. M. (1958a). The risk taken in driving under the influence of alcohol. *Brit. Med. J.* **i**, 1438-42.
- COHEN, JOHN, DEARNALEY, E. J. & HANSEL, C. E. M. (1958b). Skill and chance: variations in estimates of skill with an increasing element of chance. *Brit. J. Psychol.* **49**, 319-23.
- COHEN, JOHN & HANSEL, C. E. M. (1956a). *Risk and Gambling*. London: Longmans Green.
- COHEN, JOHN & HANSEL, C. E. M. (1956b). Experimental risk-taking. *Jahrbuch für Psychologie und Psychotherapie*, Winter, 382-8.



- COHEN, JOHN & HANSEL, C. E. M. (1957). La répartition des probabilités subjectives. *J. de Psychol.*, Jan.-March, 10-21.
- COHEN, JOHN, HANSEL, C. E. M. & DEARNALEY, E. J. (1955). The risk taken in crossing a road. *Oper. Res. Quart.* 6, 3-11.
- DEARNALEY, E. J. (1958). Delay and Risk: Influence of a delay in performance on maximum risk-taking levels. *J. Gen. Psychol.* 59, 177-83.
- POLANYI, M. (1958). *Personal Knowledge*. London: Routledge and Kegan Paul.
- VON NEUMANN, J. & MORGENSTERN, O. (1947). *Theory of Games and Economic Behaviour*, p. 28. Princeton: Princeton University Press.

(Manuscript received 28 February 1961)

## PUBLICATIONS RECENTLY RECEIVED

*Experiments in Personality.* Edited by H. J. EYSENCK. London: Routledge and Kegan Paul. 1960. Vol. I. *Psychogenetics and Psychopharmacology*. Pp. xii + 262. Vol. II. *Psychodynamics and Psychodiagnostics*. Pp. viii + 333. 40s. each.

Publication of these two volumes was designed to mark the end of the first ten years of experimental work in the Psychology Department at the Institute of Psychiatry. There is no doubt that an enormous amount of hard work has gone into those ten years. If, as a clinician, one must be disappointed that the yield of clinical applications has been relatively meagre, one has to admit that this has been largely due to research policy rather than failure to conduct the right experiments. If Eysenck and his co-workers have made it their task to discover the degree of generality in personality and behaviour, they cannot be expected to produce much that is immediately valuable for the clinician whose main interest is in individuality. Anyone interested solely in individual phenomena, therefore, is going to be disappointed by this book, as he may already have been disappointed by previous publications from the same address.

The volumes fall roughly into four parts: one, devoted to the genetics of emotional behaviour; another, an example of some good clinical research on thought disorder; a third, concerned with the theoretical link between central inhibition and excitation, Hull's reactive inhibition, and extraversion; and several further chapters, describing the statistical manipulation of the data contained in some of the others.

Thus, in one hundred well written pages, Broadhurst describes four years of work in the Maudsley animal laboratory, from its inception in 1954. The work reported concerns the genetics of emotional behaviour in the rat, emotional behaviour being defined in the conventional terms of defaecation and of locomotor activity in an open field test. In another hundred pages, in the second volume, Payne and Hewlett dispose of Goldstein's 'concreteness' hypothesis (which seems to have been a case of verbal mis-labelling), again show that retardation is inadequate as an over-all description of schizophrenic behaviour, and find that, so far as at least some schizophrenics are concerned, over-inclusive thinking is peculiar to them and not to other types of psychotics, neurotics, or normal people.

The purpose of the work described in the second parts of both volumes was: (1) To test Eysenck's drug postulates, namely, that 'depressant drugs increase cortical inhibition, decrease cortical excitation, and thereby produce extraverted behaviour patterns; stimulant drugs decrease cortical inhibition, increase cortical excitation, and thereby produce introverted behaviour patterns'. In particular, what has been tested is the effect of two depressant drugs, Doriden and meprobamate, on various aspects of behaviour. (2) To examine the relation between central processes and extraversion, employing psychiatric diagnosis or personality type as independent variable. The over-all results of these experiments, as one might expect, are not clear-cut. Some hypotheses are confirmed, others rejected. Where there is failure to confirm a prediction, this may have been due to any number of flaws, both theoretical and methodological, or to the weakness of the criterion measures of personality used in defining some of the experimental groups. There is no harm in this, as Eysenck would agree, if one regards the general framework of his theory simply as a policy for conducting research.

In the remaining chapters the Maudsley psychologists display their analytic tools; and certainly their armoury contains many that are sharp and powerful. One question in the reviewer's mind is whether the raw material on which they are used may not sometimes be a little weak. (This scarcely applies to Broadhurst's contribution.) Numbers are often small; and one wonders how representative they may be of the groups in the populations sampled. In fact, it sometimes appears in the course of the work, as in Payne's and Hewlett's, that an experimental group is not what it was first supposed to be. Such classifications as 'depressives', 'dysthymics', or 'schizophrenics' are usually pretty mixed bags. Another undoubted weakness lies in the criterion measures of personality employed. This is conceded by some of the authors; and the suggestion is made (by Willett) that, at least where one is dealing with patient material, behaviour ratings made by psychiatric staff may be more reliable than questionnaires.

To conclude: this is not a book of outstanding importance; but its publication seems to the present reviewer to be justified on a number of grounds: first, the attempt to link the manner in which learning is effected to constitutionally determined differences in personality is a worthwhile enterprise in itself, even if the information yielded is still not readily applicable in the clinic. It is something that has been largely neglected in the work of such learning-and-personality theorists as Mowrer and Millar and Dollard. Secondly, in the discussion of these experiments a lot of general psychology is introduced. It should be a salutary experience for the student, whether he is taking advanced undergraduate courses or post graduate courses in clinical psychology, to find the tools of experimental psychology being brought to bear on material that is clinically relevant. It may make clinical psychology a trifle laborious, but it may also have the effect of making experimental psychology more exciting. Thirdly, a number of unsuccessful experiments is faithfully reported and discussed. For example, neither Willett nor Holland, in two of their reports, considered themselves to have been very successful in demonstrating associations between their various measures of conditioning, learning, and perception, on the one hand, and reactive inhibition and extraversion on the other. Discussion of their lack of success laid vividly bare the enormous complexity of the variables to be controlled in experimental studies of human behaviour. One may well conclude with Willett that 'the relationship of even the simplest laboratory response to general qualities of behaviour' is as yet very far from being understood.

J. GRAHAM WHITE

*The Role of Speech in the Regulation of Normal and Abnormal Behaviour.* By A. R. LURIA. Translated by J. TIZARD. Oxford: Pergamon Press. 1961. Pp. ix + 100. 50s.

In 1958 Prof. Luria of Moscow delivered three lectures in London. Their titles were: 'The role of speech in the formation of mental processes', 'The development of the regulatory role of speech', and 'Modifications in the regulatory role of speech resulting from pathological states of the brain'. This interesting book contains the text of these lectures. Because experiments are not cited in detail and because some of the classes of pathology seem strange, this text will not be fully intelligible to readers who are hitherto unacquainted with Soviet psychology. Nevertheless, there are two important counts on which this book recommends itself. First, it outlines some of the conclusions reached by Soviet investigators about a topic which deserves more attention than it has received, that is, the functions of speech in organizing mental processes. Secondly, it cites various novel and ingeniously simple experimental situations which ought to repay further study. Prof. Luria would probably be the first to acknowledge that the utility of these situations is far from being exhausted and his brief reports of them raise issues about which the reader will want to know more. For example, a child is instructed to make a manual response whenever a light flashes on. He fails in this task. But if asked to respond vocally as well as manually, he succeeds. How does the added vocalizing exert control? Because it is vocal? Or because it produces an auditory signal so that, say, asking the child to sound a horn might do just as well? Or because there is a summation effect of making more than one response so that it would suffice, say, to ask the child to raise his foot and make the manual response? The answer does not appear to be known, but the questions seem worth asking if the role of speech is to be studied along the promising lines of this book. This is merely one of the issues suggested by the talks; and there are so many others that everyone concerned with psychological development or pathology will find stimulation in this refreshing account of Soviet work.

I. M. L. HUNTER

*Psychosomatic Aspects of Paediatrics.* Edited by RONALD MACKEITH and JOSEPH SANDLER. London: Pergamon. 1961. Pp. xi + 155. 50s.

There is an increasing realization in professional fields that complex biological and social problems require for their solution a multi-disciplinary approach. From time to time a variety of experts (sometimes from not altogether obviously related areas) are invited to spend a period of days together, pooling their resources and discussing problems of mutual interest. It is to the great credit of those responsible that, as in this case, the results of such an endeavour are often published in book form.



In 1959 a group of experts met under the auspices of the Society for Psychosomatic Research to discuss psychosomatic aspects of paediatrics; contributions were made on child psychiatry, paediatrics, learning theory, philosophy, biological research and statistics. The material is both readable and stimulating, not least in the excellent discussions which are reported verbatim.

Among outstanding contributions were a brilliantly succinct account of learning theory and personality development by Dr Hindley; the presentation of results of an ingenious investigation of recurrent pains in children by Dr Apley; and a discussion of theories of psychological development in children by Dr Bowlby. Two symposia presented novel approaches to well-worn topics: theories of psychosomatic disorders and problems of research in psychosomatics.

Nevertheless, the over-all picture in this important field is hardly encouraging; there appears to be a deficiency of firmly established facts which consequently impoverishes theory. As Dr MacKeith points out, we know very little of the full natural history of such common illnesses in childhood as asthma and epilepsy. This book thus focuses attention on an important and hitherto relatively undeveloped area. Its price, however, seems excessive.

A. D. B. CLARKE

*On Becoming a Person: A Therapist's View of Psychotherapy.* By CARL R. ROGERS. Boston: Houghton Mifflin. London: Constable. 1961. Pp. xii + 420. 37s. 6d.

This is an extraordinarily mature book. In it Rogers offers his readers a selection of his papers written between 1951 and 1961 and introduces each with his own personal appraisal of the content and intention.

Carl Rogers's psychotherapeutic technique is individual. His previous two books have given this impression and have been noteworthy for communicating Rogers's skill in verbalizing his techniques rather than for the importance of the techniques themselves. In this present collection of papers the style is highly personal starting from a chapter entitled 'This is Me' and working through by reminiscence and contemplation a wide range of professional and philosophical problems which he has encountered and examined in the course of his counselling and reflexion. The maturity derives from the contemplative attitude—at times he reminds one of a less misanthropic version of Freud.

The surprising thing is that little of his technique can be learned directly from the book. For this the earlier works must suffice. These papers reveal the development of a therapist and little of his methods, presumably because the subjectivity of the therapist is an essential of the therapeutic process. This of course condemns Rogers in the eyes of those of us who sleep better in the knowledge that we are logically positive. However, the personal struggle between the completely subjective therapist and the objective psychologist in Rogers is vividly recounted in his chapter 'Persons or Science? A Philosophical Question'. Unfortunately for this chapter some theological students 'put him on to' Kierkegaard and Buber and in consequence he tends to get bogged down slightly in some of the existential side issues. Rogers however emerges as an 'experientialist' and examines scientific method from this stand-point, loading his argument in favour of personal judgment.

The rabid psychoanalyst or anti-psychoanalyst will not find material in these papers to resolve their differences although each might select ammunition with which to carry on their struggle.

This book could be read by many psychologists of other persuasions who practice contemplation either publicly or privately. Those of us who abhor professional self-scrutiny can wait until the need arises.

R. M. MOWBRAY

*Aging and Leisure. A Research Perspective into the Meaningful Use of Time.* Ed. by ROBERT W. KLEEMEIER. New York and London: Oxford University Press. 1961. Pp. xvi + 447. 46s.

Nineteen gerontologists from the fields of Psychology and the Social Sciences have contributed to this book. The 'leisure' of the title is a term which can cover both constrained and truly free time, and contributors have interpreted it to suit themselves. The book is research orientated, and is an attempt to answer a number of questions about the activities of the elderly. Kleemeier, in his opening review, sums them up as follows '... (1) what do older people do during the non-

work time which is available to them; (2) what is the meaning or significance of the activities in which they engage during this time; and (3) how best may problems in this area be conceptualized and studied?' As Kleemeier says, the book does not give any final answers, but summarizes what is known and suggests further lines of approach.

In fact, it provides a summary of American statistical and research findings. Data are given on the amount of non-work time people have, how they get it and how they use it; and methods of obtaining factual data on the use of time are ably reviewed. It becomes clear that as yet too little is known about age differences in the use of time and that, for instance, exact figures are not available for the retired. The possible effects on the life pattern of variables such as culture, family structure, retirement and institution life, are considered in separate chapters.

Lists of activities or time charts are only one side of the picture. There are also attempts to take 'meaningfulness of activity' into account. The methods used to assess qualitative 'meaning' all seem unsatisfactory to this reviewer, but at least a start has been made on this central problem. Many of the papers discuss what makes an activity 'meaningful'. Here the writers' own values can be seen to influence the conceptual frameworks they suggest. It is particularly interesting to see that, in spite of two conferences at which contributors and invited critics discussed the papers, the writers do not all hold the same values about non-work pursuits, nor about the potential role of the aged in society. These discrepancies are themselves a helpful stimulus to thought, and these chapters are among the most valuable in the book.

Whether one is interested in research on ageing, or in practical problems among the aged, this is a book to set one thinking. It cannot help but stimulate questions and further research.

SHEILA M. CHOWN

*The Nature of Sleep.* Edited by G. E. W. WOLSTENHOLME and M. O'CONNOR. London: Churchill. 1961. Pp. xii + 416. 50s.

This is the record of a Ciba symposium under the chairmanship of Sir John Eccles. There are eighteen accounts of original work dealing with sleep and related problems. The coverage is very broad, the topics discussed range from the membrane potentials of single neurones to dreaming. The volume is a useful guide to current work and should be of value to those of any discipline interested in sleep.

The attacks that research workers have made upon the mystery of sleep have not yet breached the citadel wall although some outlying positions have been taken. The anatomical sites in the brain stem concerned with sleep are being defined with much greater exactitude than hitherto and it is certainly possible to exclude some time-honoured theories. Thus sleep is not due to anaemia of the brain, for blood flow measurements show a slight increase. Technical advances have made it possible to record from single neurones with little interference with normal activity; it has been shown that sleep is not associated with quiescence of cortical cells. Psychologists have devised measurements that reflect the deterioration of performance when sleep cannot be taken, but it is impossible at present to account for these changes in physiological terms.

The study of sleep has a long history but most of the contributors make little mention of any but their immediate predecessors; the hypothalamus is almost completely neglected. A short while ago nuclei in this region were believed to be of great importance for the genesis of sleep rhythms. The reviewer knows of no evidence that this viewpoint was wrong, but how strikingly fashions have changed! The hypothalamus is discussed nowhere except in the article on hibernation (which is excellent).

E. G. WALSH

*The Roots of Crime.* By EDWARD GLOVER. London: Imago. 1960. Pp. xiv + 422. 45s.

This second volume of Dr Glover's selected papers is, as its title proclaims, a collection of published and previously unpublished articles devoted to the subject of delinquency and criminality. To these topics are added two long essays dealing with the important problems of homosexuality and prostitution. The first paper sets the stage upon which Dr Glover portrays his understanding of these antisocial manifestations which are gathered under the concepts of criminality and delinquency. This paper, entitled 'The Roots of Crime' initiates the reader to a



developmental approach from which Dr Glover never deviates. At the same time it allows him to advance the view that psychoanalysis provides both a conceptual model and a method of treatment which the criminological psychologist can use in his work with criminals, delinquents and those who suffer from sexual deviations.

Dr Glover makes it perfectly clear that his genetic (developmental) approach to criminality must not be confused with those hypotheses which explain delinquent behaviour solely on the basis of adverse environmental experience in childhood and adolescence. The decisive factor, in his opinion, is the presence of endopsychic forces—themselves in turn a product of an interaction between constitutional factors and infantile experiences—which make their impact through the abnormal behaviour itself. In every case it is these unconscious endopsychic factors which must be uncovered and examined before the investigator can be satisfied that he has understood the reasons for the aberrant conduct.

In two essays Dr Glover explains that a complete psychological analysis of delinquent or sexually perverse behaviour can be most effectively achieved by following Freud's recommendation that psychical processes require to be described in dynamic, economic and structural terms (metapsychology). Such a study demands more than information about the unconscious conflicts (dynamic aspects) which are embodied in the symptoms or abnormal behaviour. It also requires an assessment of the economic factor—that is, the means employed whereby instinctual energies are distributed. The operation of the different unconscious defence mechanisms are described under this heading. Finally the structural situation must be investigated—especially the development and function of the super-ego. It is this psychic institution which leads to unconscious anxiety and guilt and which on this account plays an extremely important role in the genesis of certain forms of delinquency.

Metapsychological studies of this kind have a special interest for psychoanalysts. Even those, like the reviewer, who have no special experience in the field of delinquency can gain valuable knowledge which is relevant to psychoanalysis in general. A good instance of this is Dr Glover's clarification of the concept of unconscious homosexuality and its relationship to the actual perversion. The strictly psychoanalytical contributions in this volume may be found difficult by non-analytical readers but there are other articles which are discussed from a more general point of view. Representative of the latter group of topics are the essays on 'Outline of The Investigation and Treatment of Delinquency in Great Britain'; the 'Diagnosis and Treatment of Pathological Delinquency' and the chapter entitled 'Socio-legal'.

The non-analytical psychiatrist will be particularly interested in the claim that beneficial changes can be brought about in certain cases of sexual perversion by short term psychotherapy based upon psychoanalytical principles. He will also find of great practical value the criteria for the selection of cases which are likely to benefit from this form of treatment. A similar optimistic note is struck in the case of the treatment of the delinquent. Dr Glover suggests that the change in behaviour may result from the impact of the treatment team upon the delinquent individual thus leading to the appearance of healthier super-ego identifications. Another point of interest for the non-analytical psychiatrist is Dr Glover's positioning of the category of criminal psychopathy *vis-à-vis* the other mental disorders. In his opinion these conditions have closer aetiological and psychopathological associations with traumatic neuroses than with neurotic or psychotic reactions.

This book contains the fruit of Dr Glover's long experience of psychoanalysis and of delinquency. It reveals the valuable contributions which psychoanalysis can make in both research and treatment; but of further and perhaps even greater importance is the opportunity it offers the investigator of a standpoint independent of moral issues. It enables the investigator to adopt a detached attitude towards the phenomena with which he is confronted by its emphasis upon the decisive role of unconscious infantile phantasy and conflict in the genesis of antisocial and abnormal behaviour.

THOMAS FREEMAN

*Life Against Death.* By NORMAN O. BROWN. London: Routledge and Kegan Paul. 1959. Pp. xii + 366. 30s.

This is not only one of the most profound appraisals to be made of psychoanalysis as a theory of human nature, but it is also one of the most lucid expositions of Freudian psychology to appear for a long time. The author, Professor of Classics at Wesleyan University, has made a deep study



of Freud, and in this scholarly and wide-ranging book we have the rich result of a rare crossing of academic borders. For if Brown succeeds in renewing thought on the nature of culture, he succeeds no less in renewing understanding of psychoanalysis. It seems as if psychoanalysis will need periodic rediscovery, for, as Brown puts it, 'psychoanalytical terminology can be a prison house of Byzantine scholasticism in which "word-consciousness" is substituting for consciousness of the unconscious'.

Brown is right to remind us that the key concept of psychoanalysis is repression: and if the essence of the individual is repression of himself, then the essence of society is repression of the individual. In Freud's early ideas (before 1920), conflict was epitomized as that between impulses and obstructing reality. It seemed then as if repression could be ultimately explained by some objective economic necessity to work (work reality *versus* sexual pleasure). But the later Freud moved a long way from this optimistic traditionalist view of human conflict, and saw that man is the animal which represses itself and creates culture and society in order to do so. Man the social animal is thereby paradoxically also the neurotic animal. For the repressed impulse never ceases to impel to satisfaction in an endless, Faustian, self-defeating attempt to achieve 'progress'. Brown's argument is that, just as the method of psychoanalytic therapy aims to deepen the historical consciousness of the individual until he awakens from his own history as from a nightmare, so psychoanalytical consciousness as a higher stage in the general consciousness of mankind may lead man to loosen himself from the grip of the past—'to live instead of making history'. Brown examines the implications for a theory of culture of Freud's Eros and Death Instinct concepts, and of the nature of sublimation and of play, and he applies his conclusions in a series of studies of anality. Here he discusses, among other important themes, the relation between guilt and compensatory rituals of cyclical regeneration and messianic redemption; the social change which replaces reparatory giving by possessive taking and display of possessions; the dialectic of sublimation whereby the body is denied yet at the same time is secretly affirmed by means of the mechanism of projecting the repressed body into things (with some significant effects upon the development of technologies). In his final chapter on 'The Way Out', Brown asks whether man can find a new emphasis in culture, both in social life and in science, such that a greater strength of Eros might be pitted against the Death Instinct.

Most attempts to evaluate psychoanalysis for its philosophical implications have been methodological studies of the logical status of Freud's concepts. Brown's work is incomparably richer in its scope, for his concern is with the sociological and metaphysical meaning of psychoanalysis. Yet in examining these, he demonstrates, far more effectively than the logicians have succeeded in doing, the essentially dialectical nature of psychoanalytic thinking.

CECILY DE MONCHAUX

*Current Trends in Analytical Psychology.* Edited by GERHARD ADLER. London: Tavistock Publications. 1961. Pp. x + 326. 38s.

In August 1958 the first International Congress for Analytical Psychology was held at Zürich. The nineteen papers collected in this volume constitute its proceedings. The book may be recommended to psychiatrists and psychologists who wish to know more about the current practices and preoccupations of Jungian analysts; but some of it will be found difficult reading by those who are not already familiar with the Jungian point of view.

Certain papers will, however, be found particularly enlightening. Anyone who wishes to understand what Jung means by a symbol would do well to read Dr Fordham's brief paper on 'The emergence of a symbol in a five-year-old child', which is a fascinating and convincing account of how a disturbed small boy learned to deal with his aggressive impulses. Dr Hobson has attempted to apply Jungian theory to the study of a group; this paper is particularly stimulating, for it suggests new lines of research which might profitably be applied to the study of political theory and beliefs. Dr Prince and Mr Layard have both written papers on homosexuality which emphasize the fact that this so-called aberration may have a positive function in the development of the individual; and Dr Moody has some interesting observations to make upon the fact that women often lose their neurotic symptoms during pregnancy.

Every Jungian analyst will want to read this book; but the papers mentioned above will be of especial interest to those who have not been trained in this particular school, and yet want to learn more about analytical psychology. As a whole, the book testifies to the vitality of the view

of human nature with which C. G. Jung presented us. The psychologist may quail before the mass of speculation, anecdote, and unsubstantiated generalization, but it is to be hoped that some psychologists, at any rate, will set about constructing experiments on the basis of the theories here put forward. Analytical psychology is not a science; but some of Jung's hypotheses might well form the foundation of a more adequate science of the mind than any which has yet emerged.

ANTHONY STORR

*The Psychology of Abnormal Behaviour. A Dynamic Approach.* By LOUIS P. THORPE and BARNEY KATZ. 2nd edition by BARNEY KATZ and ROBERT T. LEWIS. New York: The Ronald Press. 1961. Pp. xii + 677. \$7.50.

Although written by psychologists, this is a text-book of Psychiatry for students. To the reviewer the main merits are that:

(1) In spite of the Dynamic Approach, it recognizes the importance of classification and is more logically set out than most text-books.

(2) Useful statistics are provided to orientate the student to the size and complexity of his chosen field.

(3) The experienced reader may well find that ideas which have long been in the periphery of awareness are brought into sharper focus by an apt illustration, a good diagram or a neat collecting together of such ideas. This suggests that the points will be well made for the student.

(4) The authors accept so easily the notion that Psychiatry and Psychoanalysis can both contribute to abnormal psychology that they do not need to mention it.

To the reviewer the demerits are that:

(1) Scarcely any use is made of the experimental work in clinical psychology. It surely is not that bad?

(2) Parents become such ogres that one wonders why they have not been done away with long ago. The mere fallibility of parents has become immensely boring. We need to know what particular aberrations lead to what particular sorts of breakdown and about this there is as yet little but speculation.

(3) *The authors accept a rather archaic hydraulic model of energy.*

(4) The distinction between normal and neurotic anxiety is made and then marred. In distinguishing anxiety from fear they say 'Anxiety occurs in situations where there are no signs of actual or even possible danger' (p. 144); but 'Normal anxiety manifests itself only in the face of present danger' (p. 145) and this distinguishes it from neurotic anxiety. Normal anxiety has become fear and neurotic anxiety has become anxiety, which is not presumably what was intended. Why are we never told from which anxiety we are trying to free ourselves by means of our unconscious ego-defence mechanisms? Surely it is what May has called 'the apprehensionism' 'cued off by a threat to some value which the individual holds essential to his existence as a personality'? The neurotic anxiety is bound up with the defence-mechanism.

(5) In discussing treatments no mention is made of the controlled studies which have attempted to evaluate them.

(6) Diagrams and tables are sometimes introduced without any cautionary statement about oversimplification. Thus, Table 19—1 (p. 452) shows symptoms associated with the tumours in various parts of the brain. This, without such cautions, might lead the student into an unrealistic optimism.

Although the listed demerits exceed the merits, this gives a misleading impression of what is basically an excellent book through which to get a first look at the field, whatever the reader's ultimate goal.

G. A. FOULDS

*Foundation of Psychopathology.* By JOHN C. NEMIAH. London: Oxford University Press. 1961. Pp. xx + 338. 52s.

This book has a 'folksy' quality. It assumes that psychiatry, and in particular psychopathology, have to do with human nature. Those of us psychologists and psychiatrists who take our professional concerns to be less expansive and more mundane are immediately out of sympathy with Dr Nemiah's approach. Undoubtedly his humanitarianism and personal



integrity shine through his descriptions of his contacts with patients, but the absence of a specified technical basis for his approach limits the relevance of the book.

The author states that his psychology is based on psychoanalysis and statements, whether of Proust or of his patients, are interpreted with psychoanalytic verve. His method is to present a fragment of an interview or quotation, interpret, generalize and elaborate by another example. The result is confusing as no consistent principles of psychopathology emerge, let alone foundations.

The text is intended primarily for students of medicine beginning their study of psychopathology. While it is conceivable that the book might fill a need in the more expansive American curriculum it is difficult to envisage a use for it in British medical schools. Psychology undergraduates should not be encouraged to read it as it is unrepresentative of the psychiatric approach and of case material.

R. M. MOWBRAY

*The Mentally Handicapped and their Families.* By J. TIZARD and JACQUELINE C. GRAD. (Maudsley Monographs No. 7.) London: Oxford University Press. 1961. Pp. x+146. 28s.

There can be almost no greater tragedy for a family than the birth of a severely subnormal child, and many parents have written movingly about the social, emotional and economic problems involved. Apart however, from the outstanding recent study by Saenger in New York, very little representative factual material concerning such problems has emerged. Thus the contribution by Tizard and Grad is particularly important, and, in the context of the development of new community provision for the mentally handicapped, topical.

The purpose of this survey was to investigate the problems faced by the parents of idiots and imbeciles, during both childhood and adult life, and to inquire into parental attitudes towards existing social and institutional services. Two examples, one having the child or adult at home, and the other in an institution, were studied by means of interviews and questionnaires.

A wealth of information is presented on such factors as: the cognitive, physical and temperamental status of the defectives; the economic, housing and overcrowding problems of the families, their physical and mental health and the reasons for committing children to institutions, the opinions of the parents on medical and social services and how these are used. Among the surprising facts elicited is the finding that only 12% of siblings in the home sample had mental health problems, as compared with 26% of the siblings of the institutional group. This suggests that parents with more than one difficult child are more likely to seek institutional care. The notion that the presence of a severe defective must inevitably have adverse psychological repercussions on the siblings is thus unsupported. Nevertheless, it is clear that for many parents there are grave problems, particularly economic and social.

This book gives a broad factual account of the whole problem, together with implicit and explicit recommendations for the future; it has been carried out in the best tradition of British social surveys, and is to be commended to all with interests in this field.

A. D. B. CLARKE

*The Analysis of Therapeutic Groups.* By F. KRÄUPL TAYLOR. London: Oxford University Press. 1961. Pp. viii+122. 25s.

This book is the eighth and most recent of the Maudsley Monographs. It would be worth buying for the first chapter alone, which is an erudite and well-balanced evaluation of psychotherapy. This provides a timely answer to the enthusiastic lambasting of psychotherapy which has recently become fashionable. Dr Taylor in his second chapter discusses the rationale of group psychotherapy, drawing attention to the interaction between patient and patient. This seems to make very good sense indeed. There have always been two sharply contrasted types of psychotherapeutic contact which might be termed 'professional' and 'amateur'. The former is provided by the trained and qualified expert, and the latter by the understanding and sympathetic good neighbour. In the one case, the doctor is the omniscient advisor to whom the patient goes for a cure; in the other, the good neighbour is an equal with whom the man sits down to discuss his problems which they proceed to work out together. It may be that in group therapy the best of both methods are combined.



There is a chapter in the study of therapeutic groups in which an attempt is made to quantify individual interaction within a group, and details of the statistical methods used are given in an appendix. A narrative account of a therapeutic group makes use of these various measures in describing the progress of the therapy; and there is a discussion of the results of treatment.

This is a most refreshing book which combines a sympathetic approach to psychotherapy with a sensible objective assessment of what is going on.

RALPH HETHERINGTON

*Report Writing—Psychology and Psychiatry.* By JACK J. HUBER. New York: Harper. Pp. x + 114. 21s.

Dr Huber is clearly an experienced and accomplished writer of psychological reports. It is also evident from his text that the students he supervises must be enriched by his constructive criticism and his grasp of the essentials of report writing. What is less clear is the value of this manual which in its 114 pages attempts to encompass everything from a psychiatric case history to a psychotherapeutic progress report.

A report is written for a particular reader or class of readers who must be able to understand both the language and, when it obtrudes, the theoretical orientation of the writer. This proposition is usefully elaborated by Dr Huber in several chapters which the trainee psychologist may well find helpful. Whether this formulation will enable the trainee to write better reports, is another matter. It has all the deceptive simplicity of the golf professional who tells the novice, 'The idea of the golf drive is to hit the ball three hundred yards down the fairway', and then proceeds to do so. Our post-graduate students know what is expected of them in a report; what they have to learn are the techniques necessary for achievement. In writing this book, Dr Huber assumes that it is possible to teach report writing without at the same time monitoring the process by which the psychological data are obtained, classified and evaluated. This is the assumption with which the reviewer would quarrel and out of which his dissatisfaction with the book arises.

J. DREWERY

*Statistical Reasoning in Psychology.* By R. S. RODGER. London: University Tutorial Press. Pp. viii + 204. 15s.

There are a number of trusted and tried books written on statistical method for non-mathematicians, especially for education students and psychologists. None of them is completely satisfactory, and it is never possible to recommend just one book, even if this were a desirable practice. Psychologists learning statistical method need to sample many books, to link the work with experiments in the psychological laboratory, and to have the benefit of good teaching. If those methods are adopted the field is well covered, so that any new book should preferably justify its claim to be useful by adopting some different approach. This is exactly what 'Statistical Reasoning in Psychology' attempts. It is essentially an introduction to statistical theory written from the classical view of normal variates, aiming at promoting understanding in the mind of the student, rather than giving a long list of useful techniques. Although the book is written for non-mathematicians, more of the underlying theory is covered than is usual in such books. The references for further reading are particularly good, and any student who wishes to delve deeper into the subject can easily do so. As with many books which try to achieve an understanding of mathematics without giving the whole theory, this one will be most appreciated by those whose mathematical knowledge is already considerable.

The average student of psychology will not find this an easy book in the first instance. For example, in describing the sampling distribution of Student's  $t$  an attempt is made to explain in words what a Gamma function is. Without the underlying mathematical theory this is almost impossible. And in other places the reasoning is compressed and quite difficult. But the book is one to which the student can return time after time, for further illumination, and it does give a sound background for future work. The book includes a number of very useful exercises; minor mistakes in this section are being corrected by the addition of an *errata* note, and will be eliminated in any future edition. The book makes enjoyable reading, and provided it is supplemented by good teaching will prove most useful for beginners.

JACK WRIGLEY

Gen. Psych. 53, 1

*Training for Research in Psychology. The Canadian Opinicon Conference. May 1960.*  
 Edited by KARL S. BERNHARDT. Toronto: University Press. London: Oxford University Press. 1961. Pp. x + 130. 32s.

A conference concerned with expansion and improvement of psychological research in Canada was organized by the Canadian Psychological Association to take place near Lake Opinicon, Ontario, during the last week of May 1960. The members of this conference were thirty-four senior psychologists (in university or other posts) plus seven junior psychologists (in full-time university employment). All were of Canadian nationality. Prof. Kenneth E. Clark, University of Minnesota, was the sole American invited; he presented an Evaluation Address towards the end of the conference.

Bernhardt was a senior member of the conference, and presents us with a clear, apparently adequate report of the conference proceedings, papers read, and statements issued. The conference held that legitimate methods and fields of psychological research should be defined broadly. 'Psychological research should not be restricted to the narrowly conceived methodologies or content areas which are already well established' (p. 21). Opportunities for research could best begin at the undergraduate stage. Research training for graduates should encourage 'creative thinking' within a research setting, and few formal courses (other than statistics) need be compulsory. The group expressed its awareness of the growing demand for psychologists to do applied work in various fields, such as Medicine, Business Administration, and Engineering. This demand should be met by university departments initiating or extending 'basic research' opportunities in areas relevant to applied fields. Again, 'Consideration should be given to providing University extension services for those in applied settings in order to retain and increase their active research participation and their level of professional competence' (p. 45). Academic and professional psychologists could, through co-operation, help further the development and clarification of psychological science, primarily by means of research. In his Evaluation Address, Prof. Clark stated that older disciplines, like physics and biology, had failed to hold their academic and applied workers within the 'same social structure'. Yet this conference of psychologists indicated the value of closer communication between university and non-university psychologists.

Psychologists interested in the future of British psychology may well find this book illuminating. Perhaps an 'Opinicon' conference of senior British psychologists would be worth while.

H. WALLERSTEIN

*Selected Readings on the Learning Process.* Edited by T. L. HARRIS and W. E. SCHWAHN.  
 London: Oxford University Press. 1961. Pp. xi + 428. 28s.

This compilation of readings achieves, better than most, a sense of organized diversity. It is intended for students of educational psychology, but it could help any student towards a rounded appreciation of the variety of phenomena involved in learning processes. Thirty-five papers are abstracted from published sources and arranged in eight sections. There are the expected papers on transfer of training, retention, measurement of learning, and attitudes. There are, less expectedly but none the less most welcome, papers on the problem-solving of individuals and groups, creative thinking of poets and painters, acquisition of perceptual-motor skills, and intellectual development. The selection might not suit every taste but the book's use is worth considering by anyone teaching a course on learning and problem-solving.

I. M. L. HUNTER

*Mechanism of Colour Discrimination. Proceedings of an International Symposium on the Fundamental Mechanisms of the Chromatic Discrimination in Animals and Man.* Edited by YVES GALIFRET. London and Oxford: Pergamon Press. 1960.  
 Pp. 296. £3. 3s.

This is a very important book, which gives an authoritative and up-to-date account of the latest contributions and the present state of knowledge on the mechanism of colour discrimination. The presentation of the material, which is highly technical, as experts from many sciences contributed to this symposium, is not the only factor making for difficult reading. The fact that



papers are published in their authors' original language is a mixed blessing. In a way it is quite refreshing to read about problems of colour vision in a new idiom; linguistic differences alone add to clarity in the basic understanding of the problem. Ethnical differences in the approach to the study of colour discrimination are also preserved: on the whole the French appear to pay more attention to global problems, whereas the Anglo-Saxons excel in their analytical approach. However, even those who have more than a smattering of both French and German will find the effort to understand the written word a pretty tough assignment, and very frustrating.

The main value of this book lies in communicating what was said at this International Symposium. The accounts by Von Frish and Viaud of colour discrimination in the animal world make interesting reading; they correct the notion generally held that chromatic discrimination is very seldom observed in mammals and the non-vertebrate species. The references are excellent and the table giving an extensive inventory of researches should be useful to all interested in such comparative studies.

The neurophysiological and chemical section has excellent expositions by Rushton, Galifret and Dartnall. Rushton, as usual, with almost prophetic enthusiasm, proclaims that the days of psychophysics are numbered. This is so because 'in view of the contributions from biochemistry and electrophysiology, it is no longer possible to put forward a theory of colour vision simply by the economy of the hypothesis and the beautiful symmetry of the equation to which it leads'. Galifret makes a clear exposition of what is at present known about the function of higher centres in mitigating colour perception. This section ends with a paper by Dartnall, who begins and ends his article on a rather serious note. The opening is worth quoting in full: 'It is a sobering thought that even today no one can say with complete certainty that he has ever had a cone pigment in a test tube.'

The last section deals with a topic that will be very close to the heart of all psychologists interested in vision, namely the psychophysical problem. Here we have two very valuable and readable expositions of what today are still reckoned to be the two major theories of colour vision—the Trichromatic System and the Opponent-pair Scheme. No one today seriously believes that either theory is the only way to explain the phenomenon of colour; thus, even the proponents (Stiles and Hurvich) are at pains to show the cracks that appear in their edifices. They also admit that neither theory is so exclusively formulated as not to have elements that are common to both. The eighty years or so of feud are being slowly forgotten and the differences smoothed out.

Lastly, we must mention the opening and closing sections of this Symposium. They are discussed here together because they are both speculative. Prof. Piéron in his presidential address invites us to participate in a panoramic view of breath-taking magnitude, and interesting points of interconnexions between the findings of all these varied disciplines are presented for someone to make an attempt to synthesize them.

Three such syntheses are offered—by Scheglmann, by Hansel and by Le Grand. The first attempts to reconcile the main theories of colour vision by basing its explanation on the fact that in the spectrum we perceive *five* colours. In the same manner, in the retina we find the same number of special nerve structures—the inner layers of plexiform, and these, it is claimed, could be the basic 'Farbapparate'. Hansel, on the other hand, starting from the phenomena of complementary pairs, and the resultant neutral sensation, suggests a system that would give a satisfactory explanation of all major colour anomalies. Le Grand, in his final summary on the do not stem just from the ideas of Young and Hering. The latter theory was assumed by artists as long ago as Leonardo da Vinci, while the three primaries of Young were commonplace as the nature of colour vision and have a theory that will be able to meet all the facts of visual experience. This, he interestingly suggests, could be achieved by a marriage of the two main theories into something new, which he calls the Zone Theory. Thus, protanopia and probably also tritanopia defects could be explained by lack of pigment at the cone level, while deuteranopia would have to be explained by a defect at the level of bipolars.

We may perhaps end this review with a rather nostalgic remark made by Le Grand at the end of his paper:

'J'espère, d'ailleurs, qu'il ne le sera jamais, car se serait, à mon avis, extrêmement triste, que la vision des couleurs soit définitivement expliquée.'

R. LAKOWSKI



*Brockhaus ABC der Optik.* Edited by KARL MÜTZE. Leipzig: F. A. Brockhaus Verlag. 1961. Pp. 963 and 28 plates. £3. 4s. 2d.

This book is in the best tradition of the Brockhaus Publishing Company, so well known to British readers before the war. It is a 'lexicon', which in Germany usually denotes a book in size and contents somewhere between a dictionary and a full-size encyclopaedia.

It is in fact a reference book with well-illustrated articles on some 10,000 terms used in the sciences that deal with the electromagnetic phenomenon, especially in its visual aspects. Thus, subjects such as mathematical, physical and physiological optics are dealt with side by side with such extensions of these fields as photometry, colorimetry, ophthalmology, the principles of lighting, photography, etc.

The interests of psychologists in vision are well represented; the principles of gestalt psychology, constancies, optical illusions, colour vision, dark adaptation, etc., are well expounded. In many cases references to the original sources are given, and these are international in character.

This 'lexicon' is a unique publication of a kind not found in similar form on either the British or the American book market. Topics of optics as dealt with in dictionaries or handbooks of physics usually represent only a small part of the volume and the information obtained is scanty. To get comparable articles—both in extent and in depth, specialist encyclopaedias would have to be consulted. For those who are interested only in visual aspects of physics this book is excellent value.

R. LAKOWSKI

*Vision Research. An International Journal.* Vol. 1, nos. 1/2, June 1961. Oxford: Pergamon Press. 1961. Pp. vi+201. £10 per annum (\$30.00); £3. 10s. 0d. (\$10.00) per annum if for personal use only.

The publisher and editors claim that this new journal fills a serious gap in the array of scientific journals. Up until now research workers in vision have had to publish in journals of optics, such as the *Journal of the Optical Society of America* and *Optica Acta*, or in one of the half dozen or so journals of ophthalmology. Much of the British work has gone into general scientific journals such as *Nature* and the *Proceedings of the Royal Society B*, which are not always accessible to people in other countries. This work can now be published in *Vision Research*. The journal has editors in Europe and the United States. Most of the articles will probably be in English, but summaries are given in English, French and German. At present there is no shortage of space, and technically the quality of the figures, tables and print is high.

The combined first two numbers contain Blackwell and Blackwell's description of their thirteen cases of congenital achromatopsia, which is probably the most unique study in the field of colour vision since Graham and Hsia's report on their woman with one dueteranoptic and one normal eye. This is a very good catch for the first issue of a journal, and augurs well for the future. The remaining nine research papers appear to have been hand picked to cover all the main areas of research in vision by well-known contributors. There is also a lively discussion by Wright and the late Dean Farnsworth of Pitt's isochromatic zones in protanopes and deuteranopes. It will clearly not be possible to maintain such an eminent and well-balanced set of papers in every subsequent number. We welcome this new arrival to the band of journals to which psychologists can contribute, and wish it every success.

E. C. POULTON

*Selectivity, Intuition and Halo Effects in Social Perception.* By RAGNAR ROMMETVEIT. Oslo: University Press. 1960. Pp. 173. N.Kr. 15.

Social perception, or rather the perception of people, is not one of the encouraging areas of psychological research. It is not only that decades of work in this field have provided very few generalizations from which predictions can be made. We are not even sure that we know how to disentangle the forbidding complexity of the inferential processes involved, and therefore how to ask the questions which may be essential; and often, even when we do know what question should be asked, it is by no means easy to keep it from changing beyond recognition in the transition from the legendary armchair to a situation in which the requirements for empirical testing can be reasonably satisfied.

In view of all this Prof. Rommetveit's monograph is a courageous undertaking. In the opening chapter he sets out a theoretical framework for his experiments. The bulk of the book is devoted to a description, analysis and assessment of the findings; a concluding chapter returns to theoretical issues discussed in the light of the experimental results.

The two theoretical chapters are, in this reviewer's opinion, among the best things written on the subject. Most of the basic issues are clearly stated. The general problem of social perception is treated in the framework of cognitive theory, and the basic identity of the inferential activity underlying the knowledge of both the physical and the social environment is stressed repeatedly. The argument relies heavily—and rightly so—on Brunswik's conceptual scheme, on Bruner's ideas about categorizing and on Lawrence's findings about the acquired distinctiveness of cues. This general discussion is then focused, stage by stage, on one of the essential issues: how can we transfer the subtle complexity of 'instrumental' inferences about other people from 'real life' settings to analytic laboratory situations without at the same time throwing away the baby with the bath-water?

This is the principal task of the two series of experiments conducted at the University of Minnesota and in Oslo. And it is here that one is made to think again, as in so much other writing in this field, of the uncomfortable parallel about the mountain and the mouse. Rommetveit's experiments are designed to test a number of hypotheses derived from his previous discussion. His predictions are concerned with the effects on discrimination of 'selective orientation' towards various personal attributes, with the relationship between the awareness of cues mediating information about an attribute and the subjective relevance of that attribute, with halo effects determined by subjectively dominant and unimportant attributes, etc. The experiments are ingenious and well controlled; the details of the complex experimental procedures cannot be described here. Some of the findings are statistically ambiguous, some are not easy to interpret. Rommetveit is well aware of these limitations, and it is only fair to add that they are inherent in the problem studied, and not in the methods adopted by the investigator.

The fact remains that Rommetveit has been successful in demonstrating experimentally some *systematic* differences between the 'instrumental act' of making inferences about another person's attributes and the usual laboratory judgment tasks. It is no less important that in doing this he has made a beginning in the long overdue empirical analysis of the problem of 'intuition'. The discrepancy between the ambitious theoretical introduction and the relatively modest harvest of findings shows once more that these problems are difficult, but the attempt also shows that a slow accumulation of systematic knowledge about them is not impossible.

The last chapter is in the nature of a theoretical *post-mortem*. It is devoted to a critical discussion and a synthesis of Brunswik's functionalism which, according to Rommetveit, does not pay sufficient attention to the categorizing activities of the organism, and of Kelly's functionalism, not sufficiently concerned with the stable predictable structure of the environment. Rommetveit restates here in more general terms his position that the nature of inference is the same in social perception as in other perceptual activities, and suggests the general lines of an analysis, similar to Bruner's, of active 'category attainment' in social perception. This is certainly a valuable starting-point. But the main problem is still with us: how to devise a detailed blueprint for concrete empirical work?

HENRI TAJFEL

*Social Problems in Our Time.* By S. KIRSON WEINBERG. New York: Prentice-Hall Inc. 1960. Pp. viii + 600. \$6.75.

This book aims to describe some contemporary problems of social pathology from the specifically sociological viewpoint. In Part I, after a general introduction to the study of social problems, the problems themselves are collected under two headings. In Part II deviant behaviour such as juvenile delinquency, adult crime, homosexuality, prostitution, alcoholism and personality disorders are dealt with. Part III is devoted to social problems among conventional groups such as marital conflicts, marginal age roles of adolescence and old age, and ethnic prejudice and discrimination. Each topic is organized by dividing the contributing and causal facets of the problem into three sections: (i) the social processes, (ii) the social factors, and (iii) the personality traits, the latter being deliberately put last in order to emphasize the social conditions which underlie individual perversities.

This is of course an American book about contemporary American society. This has had the



effect on a British reviewer (not so interested in the carefully presented U.S. description and statistics) of focusing attention on the theoretical analysis which the sociologist can bring to bear on social problems. And indeed the author in his preface specifically disclaims any attempt at 'an impressionistic discussion of social problems as current events'. From this theoretical standpoint the book only serves to emphasize the author's own contention on his opening page as to how little we have progressed in any real understanding of social pathology—certainly any guidelines to social change for improvement are non-existent. We have been studying juvenile delinquency as a pressing and persistent social problem for many years now and have accumulated quite a considerable knowledge about it. Is it fair to suggest that it is a measure of the inadequacy of our social diagnosis (as well as of the intractability of the problem) that the delinquency rates continue to increase?

DEREK PUGH

*Efficiency and Effort. An Analysis of Industrial Administration.* By W. BALDAMUS. London: Tavistock. 1961. Pp. viii + 139. 18s.

This book is an approach to employer-employee relations in industry, covering psychological, sociological and economic aspects. The concept *effort*, analysed into the component 'deprivations of impairment, tedium and weariness', stems from the strenuous 'physical conditions, repetitiveness and routines' of work. These feelings of deprivation respectively generate the relative satisfactions of 'inurement, traction and contentment'. Two main functions of industrial administration are to control effort stability and to increase the effort intensity of the employee. Unless wages rise (within limits) in line with increased size of effort, industrial conflict will become manifest in employee absenteeism, slowdowns and strikes. Yet '... taken as a whole, a necessary condition for successful administration is to increase the level of average effort so that this advantage to the employer is not entirely offset by a corresponding rise in wages or by provoking disturbances resulting in industrial conflict' (p. 126). It seems (to this reviewer at least) that this book overlooks an important common interest between employers and employees: maintenance and expansion of the industry itself.

H. WALLERSTEIN

*Learning, Remembering and Knowing.* By PATRICK MEREDITH. (The Teach Yourself Books.) London: English Universities Press. 1961. Pp. 174. 6s.

The 'Teach Yourself' series, now of almost encyclopaedic dimensions, already contains volumes such as 'Teach Yourself To Think', 'To Study' and 'To Live', a 'Student's Guide', and a 'Teach Yourself Personal Efficiency'. Prof. Meredith's book aims at showing the solitary student what his job is: 'to learn how to learn, to learn how to remember, and to learn how to know'. It is a sort of meta-instructional guide, twice removed from the first-order subjects of instruction.

The peculiarity of Meredith's book is that it does not just give advice or serve up empirical rules, but attempts to give the student an appreciation of the nature of the psychological processes involved in studying. Not rules nor recipes nor programmes, but an understanding of the nature of the job he is engaged in is what the student needs. So Meredith tries to get him to think fundamentally, and he does so in a refreshingly direct and vigorous way.

Not all psychologists will agree with the account of the mind that Meredith provides. Perhaps he overplays 'the schema' as a basic explanatory concept. But there is no doubt that he has written a brilliantly stimulating book—readable, suggestive, provocative, and completely free from staleness. It should waken many solitary thinkers—and perhaps even some University students—from their crocodilian torpor, and slay a number of hoary fallacies connected with the psychology of study.

L. S. HEARNshaw

*Culture in History.* Ed. by STANLEY DIAMOND. New York: Columbia University Press. 1960. Pp. xxiv + 1014. £6.

These essays in honour of Paul Radin, conceived in his lifetime but published after his death, together form a most appropriate act of homage to a man of very wide scholarship. Radin, who lectured for five years in this country under Rivers at Cambridge, was in many ways the



spiritual heir of nineteenth-century ethnologists like Sir James Frazer, but had himself a close personal knowledge of primitive communities. While he was interested in the ideas of Jung he did not consider himself a 'Jungian', but the large bibliography of his work at the end of this volume shows clearly how Radin, to the end of his life, retained his interest in the individual person within any society, and it is here that his work is of specific value to the psychologist.

Distinguished contributors abound: Redfield, Hallowell, White, Kluckhohn, Lowie, Loeb, Evans-Pritchard, Levi-Strauss, Tillich, Kroeber and many others. This very wide range of talents makes the task of the reviewer exceedingly hard. The fifty-two essays cover too many disparate fields to be adequately discussed in a short review. They are grouped under five headings, but these tend to mislead as the range within each is so great. Thus we have: I. The Primitive World View; II. Approaches to Culture (which includes such differing topics as 'The native dog in the Polynesian system of values' and 'Reflexions on the ontology of rice'); III. Ritual, Religion and Myth ('The dialectic of Christianity' and 'Remarques sur le Rope Trick'); IV. History, Social Theory and Law; and V. Language ('Language, evolution and purposive behaviour' and 'Some genetic affiliations of Algonkian').

The ethnography is of interest, and light is thrown on the Ojibwa, Navaho, Winnebago, Shoshoni, Incas, Yurok and other peoples. To the psychologist the most valuable aspect of all these contributions is that the great majority of the scholars concerned have been at pains to deal with the thinking and 'projective systems' of the various cultures investigated. There is, for example, a most interesting article by Redfield on 'Thinker and intellectual in primitive society', and Kurt Goldstein's contribution provides a most useful antidote to the assumption that the thinking ability of primitive peoples is in some definite way 'inferior'.

The main ethos of the collection, if one indeed can be found, lies in the attempts to bridge methodological and theoretical gulfs. This is again a tribute to the inspirer of the collection. It is very refreshing to read anthropologists who are not afraid to discuss the values of human beings and the motivational categories which, in the opinion of this reviewer, are an integral part of the explanation of social behaviour.

In brief, while it is impossible to recommend this book as in any way a standard text, it will nevertheless repay any reader (or browser) with many stimuli to thought. It is a magnificent bedside book for a long and cold winter.

S. G. LEE

*Educating Gifted Children.* By ROBERT F. DEHAAN and ROBERT J. HAVIGHURST.  
London: The University of Chicago Press. 1961. Pp. x+362. 42s.

This 'revised and enlarged' edition of a book first published in 1957 offers some useful new material on a subject which has enjoyed increasing attention in America in recent years; and the book is now more coherently and solidly structured. Probably its greatest attractions are in the earlier chapters which deal with the need to discover gifted children, and methods of doing so, and demonstrate losses of talent in groups under-privileged by reason of social class, race, sex. There is also a good systematic discussion of the methods of enrichment, grouping and acceleration; it is pleasant to find support for the acceleration policy, which is usually over-harshly judged.

Some weaknesses arise from the attempt to be comprehensive. Certainly artistic genius and high social ability are important forms of giftedness; but the impact of the book would be greater if it concentrated on the form about which the authors have really most to say—intellectual ability—and gave more than the present meagre allocation of space to the extremely gifted child. This would happily avoid the conscientious excursions which do not really give much guidance towards developing artistic abilities and 'social leadership'. Indeed the discussion of the latter reveals some conflict between duly reported research findings which fail to provide a distinctive pattern of leadership characteristics and the author's belief that nevertheless something should be done about discovering and educating future leaders. The ultimate state of bafflement may be illustrated by one of the questions set at the end of this chapter: 'What abilities will future leaders need in view of the unpredictability of the future?' Similarly, the discussions of motivation and 'creativity' could be undertaken at a deeper and more critical level. The attempt to devise classroom exercises (e.g. 'brainstorming' and fluency practice) to develop 'creativity' and originality is gallant but unconvincing.

The book, however, is intended for a wide public, including interested laymen as well as

students and teachers. It should be of value in making them aware of the many aspects—psychological, educational and social—of developing ability. It can also be extremely interesting when read from another point of view: that of studying some of the social and philosophical principles underlying the American system of education.

MARGARET B. SUTHERLAND

*Examinations and English Education.* Edited by S. WISEMAN. Manchester: Manchester University Press. 1961. Pp. xx+188. 21s.

It is an excellent custom of Manchester University School of Education to hold annually a series of lectures on educational topics. This volume is based on a series about examinations. The seven chapters cover the historical developments of English school examinations, their relation to Grammar, Secondary Modern and Primary Schools, their efficiency, with particular attention to selection, and the questions arising out of attempts to examine 'general education'. The interests of the writers vary, the administrative, the educational and the psychological all being represented. In each a workman-like presentation has been achieved.

The '11-plus' is discussed, in a commendable objective way, but is not allowed to dominate. The main centre of interest is rather in examinations at school leaving and University entrance level. In such a wide survey, certain shortcomings are unavoidable. Some of the writers seem to find themselves cramped for space in dealing with the ramifications of the English examination system; the uninitiated may occasionally be confused among the G.C.E.s, the J.M.B.s, the S.S.E.C.s, etc.; and the major question, the effects of examinations on the manner and content of school education, is discussed in terms of hopes and opinions—in any case, solid information is just not available. Nevertheless, these lectures contain much of value and interest, and fully justify their publication for a wider audience.

JAMES MAXWELL

*The Management of Innovation.* By TOM BURNS and G. M. STALKER. London: Tavistock. 1961. Pp. 269. 30s.

This book is concerned with problems that face companies or industries when the development of new processes becomes unavoidable. It is not so much, therefore, a study of adaptation to change as of adaptation to the need to change. The work is a study of management—management relationships with all the inherent problems, status and otherwise, arising between traditional managers and technical incomers, some of whom may not have been in business before. Attempts to set up the technicians in research laboratories physically isolated from production are seen partly as the wish of the business not to link research with production, and partly as the wish of the technicians to keep themselves 'academically' apart from production. The industrial zone studied in some detail is the electronics industry, a prototype of newer forms of industrial enterprise with an increasingly high technical component. The book covers twenty firms and gives some fascinating pictures of ways in which companies respond to the need to incorporate this new technical skill into their organization. Many of these firms had up until this point been cushioned from harsh commercial reality by being committed almost completely to Government contract work. Research became a serious word for some of them for the first time.

Two types of organization, mechanistic and organic, have been identified by the writers—static and dynamic might equally well apply. Sets of recognition criteria are provided (pages 119–23) but the main difference is simple. The mechanistic form is appropriate to static conditions—seen requirements for action, which cannot be broken down or distributed automatically, arising from the functional roles defined within a hierarchic structure.

Case history material obtained in interview from members of management is used in a somewhat anecdotal way to support the argument. It seems in places to touch on problems inside the factory—of relationships, status, rivalries—that were unlikely to have been brought into the open beforehand. It may have been that the reading of the draft of the book was the first time that certain managers were aware of how they or certain of their actions were perceived by their colleagues. It does not seem enough to say 'there are no interpretations and appraisals contained in any part of this report which have not been communicated at some time or other to persons involved in the situations at issue'. While information collected was intended to be used in a



general context, the researcher had an effect on the individual firm and one would want to know what happened as a result. A follow-up study now of all firms involved might prove very valuable, since this work ended in 1958.

Another regret that might be expressed is that the studies were completely management-directed. The covering of some component of worker attitude to the need to adapt would have been valuable, since whether a system be organic or mechanistic is of considerable significance not only to management but to workers. The bottlenecks, the rivalries, the artificial boundaries, the tactics used by certain members of management against other managers impinge in many instances most powerfully and painfully on the workers on the floor.

The important contribution of this book is that the lessons from it have relevance in a much wider context than the electronics sector of British industry. The integration of specialists of all kinds, in the field of human relations, for example, into working communities, and the optimum utilization of their contribution is becoming acute for more and more industries.

A. MITCHELL

*Abnormal Psychology*. By W. J. COLVILLE, T. W. COSTELLO and F. L. ROUKE. London: Constable. 1961. Pp. xvi+298. 14s.

*Jordi*. By T. I. RUBIN. New York: Macmillan. 1960. Pp. vi+73. 10s.

*Better and Better Every Day*. By E. COUÉ and C. H. BROOKS. London: Unwin Books. 1960. Pp. 158. 6s.

*Nature Hits Back*. By M. LAWRIE. Kingswood, Surrey: World's Work Ltd. 1960. Pp. 190. 3s. 6d.

*Graphology*. By H. A. RAND. Cambridge, Mass.: Sci-Art. 1961. Pp. 208. \$1.90.

*Intelligence: Its Evolution and Forms*. By G. VIAUD. (Translated from the French by A. J. POMERANS.) London: Hutchinson. 1960. Pp. 127. 5s.

*Paths of Loneliness*. By MARGARET M. WOOD. London: Oxford University Press. 1961. Pp. xii+237. 12s.

*Journey Through Adolescence*. By DORIS ODLUM. Harmondsworth, Middlesex: Penguin. 1961. Pp. 160. 2s. 6d.

The above paperbacks cover a wide range of human behaviour and experience over the areas of cognition, emotion and motivation. The first four focus on diagnosis and treatment of psychological disorders. Handwriting and intelligence are dealt with by the next two books respectively; and the final two cover problems of adjustment and development faced by normal people.

*Abnormal Psychology* is a fairly thorough and clear text-book with comprehensive coverage of psychological factors underlying abnormal behaviour. Physiological and cultural factors appear less well covered; however, a fair number of books are here referred to (pp. ix-xv). This summary of information about abnormal behaviour during the presence (or absence) of brain tissue damage should be an excellent guide for interested students as well as a valuable reference for clinical psychologists. The next book, *Jordi*, follows a schizophrenic boy through four years of treatment in a psychiatric day school. Even though this child is fictional, his life from age eight to twelve is based on Rubin's factual experience as a practising psychiatrist. At the age of twelve Jordi, though still in need of psychiatric treatment, has improved enough for transfer to an ordinary day school. The author succeeds in conveying the child's subjective panic and suffering while giving us room for objective appreciation of his development. That almost all suffering may be cured by the ailing individual himself is the thesis of *Better and Better Every Day*. These cures, it is claimed, can be effected by means of Coué's method of auto-suggestion (first published in 1922), which involves use of imagination or thought without the interference of will or effort. The only persons not to profit by this method are said to be the mentally deficient or 'those who are unwilling to understand' (p. 16). Another panacea for most illnesses is discussed in *Nature Hits Back* (a revision of a 1936 publication). Good natural



food, fresh air and water, with abundant exercise would prevent chronic oversensitivity, despondency and fatigue; mildly incapacitating (yet tragic), these disorders are very prevalent in this country. Lack of good nutrition here is shown by prevalence of tooth decay, suffered by 97 % of the population. Though peripheral to his essay on nutrition, the author's chapter on sex (pp. 137-166) is remarkably enlightened.

Coming into the area of more normal behaviour and experience, we have a handbook on *Graphology*; a reference for analysing of handwriting. A well-organized system for relating graphic signs to personality traits is presented but the system's validity appears dubious. Demonstration of trait deviations in terms of handwriting deviations from '...the copybook model letters...' (p. 36) probably must await more adequate understanding of personality dynamics. The author of *Intelligence* distinguishes between 'practical' and 'speculative' intelligence. The former is tied to perception of immediately given situations, whereas with speculative intelligence '...adaptations of ideas precede adaptations of movements' (p. 72). Only adult humans possess speculative (i.e. conceptual, logical, rational) intelligence; the relevance of speech is here indicated.

*Paths of Loneliness* describes the forces in complex modern society which lead towards separation between groups and often result in isolation between individuals. Military occupations, social strata, and criminal gangs make for group separation; old age, unemployment, and physical disability may produce individual isolation. Reactions to these 'sundering powers' range from egotistic delinquency to successful attempts for social participation. With the last book, *Journey Through Adolescence*, we ride through the years 11-17, i.e. adolescence, a stage of tempestuous physical and mental growth. This book gives a very adequate description of adolescent responses to home, school, work, community. Also discussed are the joys and problems of adolescent friendship and love. When aided by understanding social guidance normal development is likely to occur, resulting in a mature integration of the young adult with the surrounding environment. The last two books, covering normal adjustment and development, prove most refreshing reading.

H. WALLERSTEIN

## THE ABSOLUTE THRESHOLD OF ELECTRIC SHOCK

By ROBERT T. GREEN

*University College London*

Thresholds for electric shock were obtained under twelve conditions. Three types of stimulus were used; constant voltage, constant current, and constant power. These, combined with three sizes of electrode, gave nine conditions in all. A further three thresholds were obtained using the three types of stimulus and the medium size electrodes, with the addition of electrode jelly to improve contact.

The results indicated that power is the relevant variable, and that sensitivity to threshold shocks probably depends on the amount of power dissipated in the immediate vicinity of the skin receptors. This view is supported by the observation that skin thickness seems to be of considerable importance—the thicker the skin, the higher the threshold.

The influence of ambient temperature and humidity were found to be inoperative within the limits investigated. A low skin temperature did, however, raise the threshold.

Subjective sensations associated with electric shock are discussed, including an anomalous lowering of the threshold as a result of rapidly repeated strong stimulation.

Implications for the use of electric shock as a negative reinforcer in psychological experiments and as an agent in 'therapy' are considered. In both cases there would appear to be sound arguments for using a constant wattage a.c. stimulus to counteract the variations in the impedance of a biological circuit.

### PROBLEM

Whereas thresholds, both absolute and differential, have long been established in many sensory modalities, few investigations have been concerned with thresholds of electric shock. This can hardly be due to the 'pure' nature of the problem, since it is no more recondite than those classics of psychophysics, judgements of weights or the two-point tactual threshold. In fact, since shock is used widely both as a negative reinforcer in experiments and as an agent in shock 'therapy' the need for some basic information regarding electric shock as a stimulus is hard to gainsay.

More probably, the dearth of experiments in this modality is related to the problem of defining the relevant parameters. With weights and distances it is clear what dimensions are involved, but with electric shock the number of possible variables influencing thresholds is manifold. Some of the more obvious ones may be listed as follows:

Alternating *v.* direct current; waveform (e.g. square *v.* sawtooth); duration of shock; onset *v.* termination of shock; anode *v.* cathode; voltage; amperage; wattage; voltage density; current density; wattage density; size of electrode; shape of electrode; material of electrode; pressure of contact; nature of contact (e.g. dry *v.* wet); locus on surface of body; temperature; humidity; individual differences; sequential effects.

Formidable though this list may appear, a start has to be made somewhere. There is little point in controlling voltage or amperage in our experiments, or 'therapy', if these happen not to be the relevant parameters. Earlier work (Forbes & Bernstein, 1935; Hill *et al.* 1952; Green, 1958) has already indicated that if any of these variables is to be accorded greater importance than the others, then it is the wattage that should be controlled. Perhaps this is a contributory reason for previous lack of progress in this area. A constant wattage stimulator has never been available, and presents difficulties in electronic design.

In the present experiments the problem has been delimited by using a rectangular direct current stimulus of 1 sec. duration. Three types of stimulus were used: constant voltage, constant current, and constant power. In combination with these, three sizes of electrode were used, making nine conditions in all. The electrodes were of silver steel and circular in shape. The smallest were 0.075 in. diameter, the middle pair were 0.15 in., and the largest 0.3 in. The loci used were the ball of thumb and index finger of the left hand, the positive electrode being on the finger and the earth on the thumb. Pilot runs indicated that, surprisingly enough, the pressure of contact made little difference, within wide limits, to thresholds. Provided pressure is sufficient to ensure complete contact, and not so great as to produce rapid numbing of the skin receptors, no further control of this parameter appears necessary. More careful observation may modify this view.

In addition to the nine conditions mentioned, an electrode jelly was used with the medium size electrodes for the three types of stimulus, giving another three thresholds. The reason for this is discussed later.

Temperature and humidity were not controlled directly, but were noted, so that their influence, if any, could be gauged.

Individual differences are, of course, always with us, and in this case the indications are that wide differences do exist. The present study, however, is not concerned with defining the distribution of these differences nor with providing norms.

#### APPARATUS\*

Since constant current and constant voltage circuits are commonplace they require no description. The constant wattage system may be briefly described. The fundamental difficulties to be overcome derive from the fact that the reactance of living tissue is not precisely known and may vary quite rapidly with time. If the power dissipated is to be maintained at a constant value it is necessary, in effect, to monitor continuously the instantaneous product of voltage and current, to compare this with a reference, and use the resultant error signal to control the output of the generator.

This is achieved by continuously comparing the load† current with a function proportional to the reciprocal of the load voltage.‡ This function is derived by width-modulation, followed by integration and smoothing, of pulses generated by a phantastron triggered at 2.5 kecy./sec. The error signal is arranged to control the generator output by way of a high-gain d.c. amplifier.

Simple switching allows the same instrument to be used to generate constant voltage and constant current stimuli.

#### PROCEDURE

Some 11 hr. of training preceded the main experiment. This was necessary for several reasons. First, rough thresholds and suitable intervals for the variables of voltage, amperage, and wattage had to be established. Secondly, a satisfactory procedure had to be developed, particularly with regard to serial presentation of the stimuli and collection of data. Thirdly, it required considerable sophistication on the subject's part (in this case the author, referred to as  $S_E$ ) to detect a threshold shock. At threshold values the distinction between shock and no-shock is not clear-cut, and it takes some time to learn to respond to the relevant cues. Statistical

\* This was designed by Mr B. Corbett and supplied by Messrs System Research Ltd.

† Although when using direct current it is usual to equate load with resistance, the fact that the stimulus is felt only on the make or break indicates that factors such as capacitance, inductance, and back e.m.f. may also be involved. See Curtis (1950).

‡  $P = IV$ , so  $I = P/V$ . If  $P$  is to be constant,  $I$  must therefore be proportional to the reciprocal of  $V$ .



theories of perception, such as those put forward by Solomons (1900), Tanner & Swets (1954), and Gregory (1956), seem very apposite in this context. In the earlier stages of training, the 'confidence' level fluctuated, and the 'doubtful' category was used frequently around the middle of the range of uncertainty. With practice, the 'confidence' level hardened, so that Type I errors would have been minimized had false stimuli been used. This reluctance to run a risk of accepting false positives became apparent when the electrodes were accidentally disconnected from the stimulator during a practice session. Some 140 false stimuli were presented, all of which drew a negative response, much to  $S_E$ 's surprise when the higher values were being set and displayed on the stimulator. It was this discrepancy between the displayed values and lack of subjective shock that led to the tracing of the fault.

The procedure eventually adopted was scheduled to take about 2 hr. per daily session. While the apparatus was warming up, the temperature and humidity were noted, and  $S_E$ 's hands were washed in warm water, rinsed and dried. The electrodes were arranged so that  $S_E$  could fit them comfortably between the index finger and thumb of his left hand without fatigue.

All nine combinations of type of stimulus and electrode size were used on each day. The programme called for a minimum of seventy stimuli under each condition, ten stimuli for each of seven selected values, making a total of at least 630 judgements per day. The order of presentation of conditions, and of stimuli within each condition, was randomized every 2 days. On the second day of any pair of days the sequence adopted was the reverse of that used on the previous day. In this way practice, fatigue, and other sequential effects were distributed as fairly as possible.

Three categories of judgement were allowed: *stimulus detected*, *doubtful*, *stimulus undetected*. Where the middle category was used, that particular trial was discarded and replaced, so that a total of ten positive and negative judgements was obtained. In this way the disadvantages of using a forced judgement were avoided, while the problems posed by the middle category in the analysis were also circumvented. In the event, the middle category was hardly ever needed during the experiment proper. After each set of ten judgements had been obtained,  $S_E$  removed his finger and thumb from the electrodes and rested the left hand.

On a particular day it sometimes transpired that the seven values chosen to cover the range for a given condition were inappropriate or insufficient. In these cases the values were rearranged so that the complete range of uncertainty was covered as far as possible. This sometimes meant using considerably more than the seven values intended.

This procedure was carried out for 10 days to furnish a total of 100 judgements for each value of the variable under all conditions, thus providing a minimum of 700 judgements upon which to calculate each threshold.

After the data for the nine conditions had been studied it was decided to repeat the experiment with the medium size electrodes, using electrode jelly to improve contact. The procedure was similar to that already described, except that it was possible to employ the reversed pattern of stimuli presentation on the same day instead of on successive days, giving twenty judgements per variable value per condition per day. Fresh jelly was applied to the thumb and finger immediately prior to each run of ten judgements.

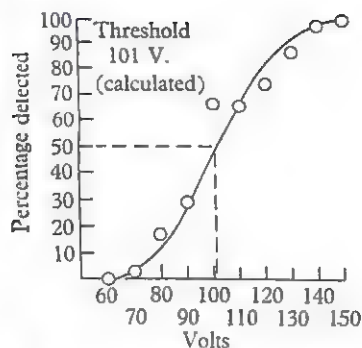


Fig. 1. Large dry electrodes—constant voltage.

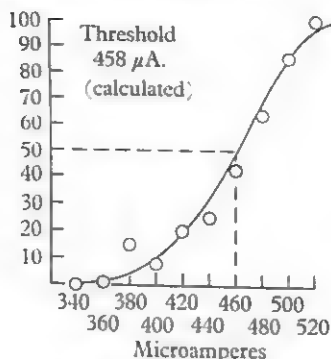


Fig. 2. Large dry electrodes—constant current.

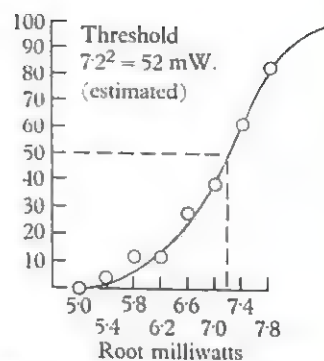


Fig. 3. Large dry electrodes—constant power.

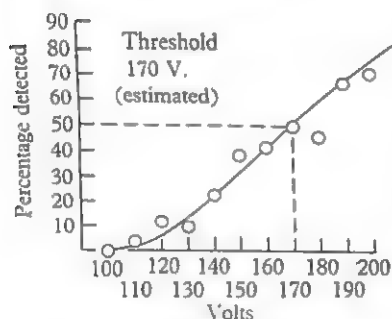


Fig. 4. Medium dry electrodes—constant voltage.

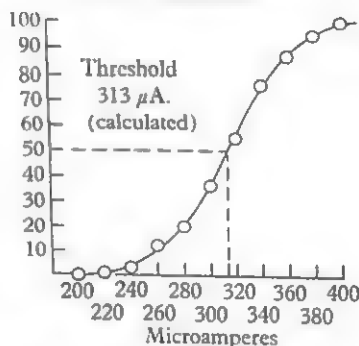


Fig. 5. Medium dry electrodes—constant current.

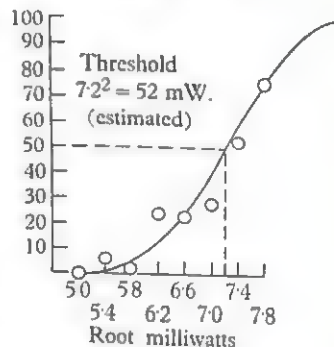


Fig. 6. Medium dry electrodes—constant power.

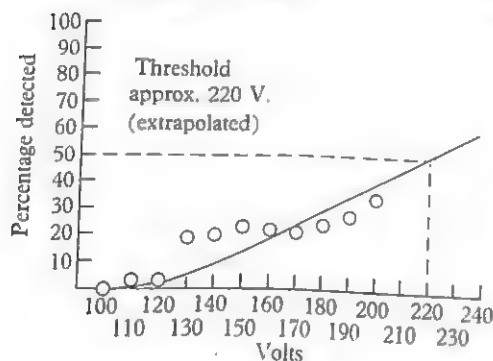


Fig. 7. Small dry electrodes—constant voltage.

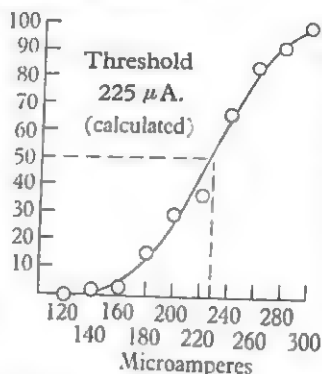


Fig. 8. Small dry electrodes—constant current.

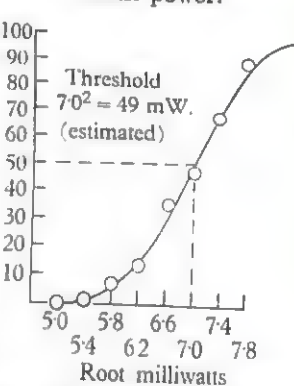


Fig. 9. Small dry electrodes—constant power.

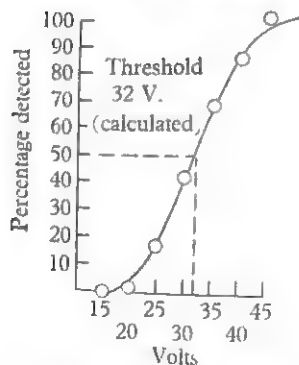


Fig. 10. Wet electrodes—constant voltage.

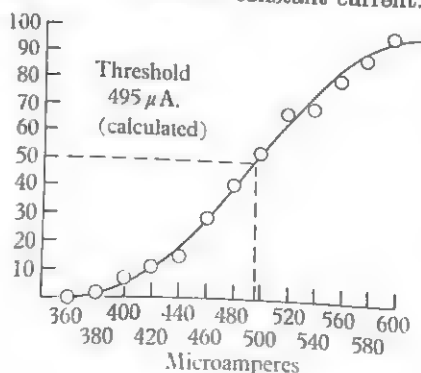


Fig. 11. Wet electrodes—constant current.

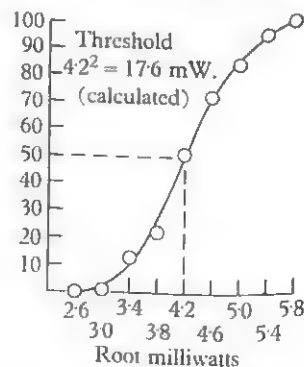


Fig. 12. Wet electrodes—constant power.

## RESULTS

Graphs were drawn for all twelve conditions and a free-hand ogive fitted. In seven of these twelve cases it was also possible to calculate the threshold directly from the data, independently of the graph.

Regrettably, in making out the specification for the apparatus, it had not been foreseen that stimuli in excess of 200 V. or 60 mW. would be needed. Since a mathematical analysis requires that the complete range of uncertainty be covered, estimated values of the threshold had to be used where the data were incomplete. Fortunately, the results were significantly clear-cut to make this shortcoming of small consequence.

With dry electrodes the threshold in terms of power remains very nearly constant at about 50 mW.,\* regardless of the area of contact. The current threshold, however, increases from 225  $\mu$ A. for the smallest electrodes to 458  $\mu$ A. for the largest pair. The threshold in terms of current density is therefore inversely related to the size of the electrode. The voltage threshold rises from 101 V. with the largest electrodes to about 220 V. with the smallest.

Clearly, then, power is the single best measure of the shock threshold, since a 16-fold increase in area of contact does not materially alter this index, but does affect other indices substantially and systematically.

This statement, however, demands qualification in the light of the results obtained when electrode jelly was used. Under these conditions the threshold drops to 18 mW. Evidently, it is not the total power as such that counts, but *where* it is dissipated. With dry contacts much of the power is spent overcoming resistance at the boundary of contact between the electrodes and the skin. With electrode jelly this source of resistance is much reduced and far less power is needed to produce a detectable stimulus.

This suggests that the crucial factor is the power dissipated in the subcutaneous tissues where the receptors actually lie. This view is supported by the casual observation that there would appear to be a simple, direct relationship between the thickness of a subject's skin and his threshold. A fine, delicate skin, such as many women possess, seems to go with a low threshold, while the man who works with his hands and has a coarse, heavy skin has a far higher threshold. In the case of  $S_E$  this effect was most noticeable. After a week-end of manual labour (handling bricks), sensitivity to shock was reduced markedly, and it took several days of non-manual activity before the threshold began to drop again. It was due to this unforeseen influence more than anything else that the apparatus proved unable to cope with the higher voltages thus made necessary.

The influence of temperature was studied by relating the number of positive responses on each day to the temperature. No consistent trend was apparent, but since the temperature ranged only between 66.5° and 71° F. over the period concerned it would not be safe to assume that no relationship exists. With a second (female)

\* Gilmer (1937) reported a remarkably similar result using 2 mm. dia. brass electrodes on the volar surfaces of the index or middle fingers. The average threshold of four subjects was around 50 mW. for an a.c. stimulus of 500–1000 cyc./sec. Higher frequencies produced higher thresholds, while lower frequencies produced a fairly stable threshold around 35 mW.



subject ( $S_2$ ) a temperature range of 68°–75° F. again revealed no consistent trend. *A priori* one would expect, since the importance of contact resistance has been demonstrated, that on warmer days the tendency to sweat would be greater and hence the threshold would be lower.

Temperature *was* important in another direction. Although room temperature seemed of little consequence in the range referred to, skin temperature had to be watched carefully. During one practice period, when  $S_E$  had come directly into the experimental situation from a cool morning, his threshold was much higher. Half an hour later, when his skin temperature had stabilized in the higher room temperature, the threshold was back to normal. A cool skin would thus seem to indicate that the receptors involved are slightly numb, and therefore less sensitive.

Serial effects were assessed in a similar manner. During the first 5 days 40% of the stimuli were detected, against 29% during the latter 5 days. This suggests strongly that sensitivity decreased during the course of the experiment. It is almost certain, however, that this was due, not to fatigue, but to the effect of manual labour and the consequent hardening and/or thickening of the skin already referred to. No serial effects were detected with  $S_2$ .

Relative humidity did appear to affect the threshold in a consistent way with  $S_E$ , but this finding was not confirmed with  $S_2$ . The apparent relationship between humidity and sensitivity turned out to be an artifact of the serial effect. It so happened that the changes in sensitivity due to skin thickening roughly coincided with changes in humidity.

Since an experimenter being his own subject invites the classic bias of stimulus error, it was thought advisable to repeat the main part of the experiment with a completely naïve subject. The three types of stimulus were used with the largest and smallest pairs of electrodes, without electrode jelly.

The results are summarized in Table 1.

Table 1. *Shock threshold for  $S_2$  (series 1)*

	Constant power (mW.)	Constant current ( $\mu$ A.)	Constant voltage (V.)	Product $IV$ (mW.)
Large electrodes	32.2	362	44	15.9
Small electrodes	31.9	198	121	24.0

Table 2. *Shock threshold for  $S_2$  (series 2)*

	Constant power (mW.)	Constant current ( $\mu$ A.)	Constant voltage (V.)	Product $IV$ (mW.)
Large electrodes	28.7	334	43	14.4
Small electrodes	30.2	238	117	27.8

Under constant power conditions the threshold was close to 32 mW. with both large and small electrodes. The current threshold increased substantially with the larger electrodes, while the voltage threshold dropped proportionally even more. This result therefore confirms the main finding that under constant power conditions the threshold in milliwatts is independent of the area of contact.

The one disturbing feature is that the product  $IV$  does not cross-check with the constant power threshold. In the original experiment this cross-check was fairly

close. To settle the question the apparatus was tested, serviced, and recalibrated, and the whole experiment repeated with  $S_2$ . The results, given in Table 2, were closely similar to the previous outcome, including the anomalous failure to cross-check.

At first glance such a discrepancy seems highly improbable. If the threshold described in power terms is 32 mW., then this same threshold must be describable in terms of a current and voltage whose product is this value. We can escape this dilemma only by recognizing that if we choose to control a given parameter of the stimulus we obtain a threshold specific to those conditions, and that this figure will not necessarily agree with the threshold obtained when some other parameter is monitored.

The issue may be analysed in more concrete terms. Under constant power conditions a stimulus is delivered that is rectangular with respect to power. If, however, the impedance of the load varies, as indeed it does with a biological circuit, then the stimulus in current or voltage terms will be anything but rectangular. Under constant voltage and constant current conditions the same argument holds *mutatis mutandis*, so that a stimulus that is rectangular with respect to power under constant power conditions will be fluctuating in some uncontrolled manner with respect to power under constant voltage or constant current conditions.

Thus, if power is the relevant variable, under constant voltage or constant current conditions we can talk only of the *average* power dissipated. Actual power will differ from the average value as the impedance varies, and will peak in excess of this average value. Since, as the on-off phenomenon shows, we are sensitive to change rather than steady state stimuli, it follows that the product  $IV$  will be less than the power threshold obtained under constant power conditions, provided that the subject's impedance varies sufficiently to distort the shape of the power wave under constant voltage and/or constant current conditions. The discrepancy between the product  $IV$  and threshold obtained under constant power conditions is therefore an index of the degree to which the subject's impedance varies.

#### DISCUSSION

The sensations involved in detecting threshold shocks are of some interest. The weakest stimuli that can be detected feel more like a gentle pressure from the area of contact than anything else. This is felt more readily on the earth electrode when the circuit is made, and on the positive electrode, more weakly, when the circuit is broken. If a slightly stronger stimulus is used, these pressure sensations are felt on both electrodes on make and break. As the wattage is increased, a pricking sensation occurs which eventually becomes sharp enough to mask the pressure sensation entirely. If this sharper type of stimulus is repeated fairly rapidly, say once every 3 or 4 sec., then the subjective intensity increases progressively until definite pain is felt, sufficient to discourage further trials at that level. If the procedure is continued past this point a burning sensation of some intensity ensues. On cessation of the stimulus the feeling of 'bright' heat moderates to a warm 'glow' that disappears after a second or two.

The increased sensitivity, produced by repeated application of a sharp stimulus, lasts for perhaps a minute or more, so that a weaker stimulus that was previously difficult to detect comes well above threshold. It is as if the receptors involved had become 'tender', and hence more sensitive, as a result of the stronger stimulation.

This finding appears to be in conflict with the classical work by Adrian (1928) on the adaptation of sense organs to an electrical stimulus. At the same time, the fact that the stimulus is felt only on the make or break is directly in line with Adrian's principles. An interesting aspect of this phenomenon is that it can be obtained under either constant current or constant voltage conditions, although it is perhaps more noticeable with a constant voltage stimulus. (It was not possible to obtain a sufficiently strong stimulus from the constant power output to complete the comparison.) This suggests that the changes involved occur in the receptors or centrally, and are not simply due to a reduced reactance of the whole biological circuit. Under constant current conditions a net reduction of reactance would lead to lower voltages and less power.

However, the question of where this power is dissipated is again relevant. It might be supposed that changes take place in the receptor tissues so that the reactance of this part of the biological circuit is increased relative to the skin and other components. If this were so, then under all three stimulation conditions more power would be dissipated at the receptors and less elsewhere, at least during onset and cessation of shock. It is known that capacitative changes take place in active nerve tissue. Alternatively, the changes that take place at the receptor may be such that, although the reactance is not increased, the tissue responds more readily to a given stimulus. This seems rather more plausible in view of the fact that 'excitation is accompanied by a measurable reduction of the nerve impedance' (Katz, 1939, p. 80).

Parallel with these changes in the skin sensations another sensation appears. If a stimulus slightly stronger than that which begins to produce a sharp pricking sensation is used, then the shock is no longer localized to the area of contact. Sensations of a different character occur in the musculature of the finger and thumb. There is a sort of twitching sensation, as if the muscles involved were going into involuntary contraction, which is probably the case. Since both flexors and extensors are involved, there is also a mild sensation of momentary paralysis. This sensation occurs only as the circuit is completed. It is not associated with the continuation or cessation of the shock. It is not in itself painful.

It is also worth noting that the pain threshold is not very far beyond the region of uncertainty. With the large electrodes, for instance,  $S_E$  usually found a shock of 140 V. distinctly unpleasant, while  $S_2$  was similarly affected by 70 V.

The role of attention is also of importance at threshold values. Concentration of attention on the hand increases sensitivity, so that with values of shock close to the middle of the range of uncertainty the subject can in effect 'choose' to feel the stimulus or not, as he pleases. Letting the eyes go out of focus or, better still, closing them assists this operation. At these values it is therefore essential to maintain as constant a degree of attention as possible to obtain reliable results. During the experiment this was done by concentrating attention on the hand, with the eyes open, but out of focus. An external noise, such as aircraft passing low overhead, reduces sensitivity. On the other hand, a sudden loud noise, such as the ringing of a telephone, lowers the threshold considerably. It is presumably a question of whether the noise is sufficiently startling to produce a sweating response. Whenever such external distractions occurred, which fortunately was not often, the collection of data was suspended.

It will be noted that in plotting the curves for the constant power condition the values are expressed in terms of the square root of the applied power. Had the



'straight' values been used the curves would have lost their symmetry and extended logarithmically along the abscissa. The logic of this is easy to follow if we remember that power is the product of voltage and current. For a given resistance the power dissipated is directly proportional to the square of the applied voltage. Or, put another way, doubling the voltage doubles the current, and hence increases the power fourfold. So, if we wish the power co-ordinate to have the same dimensions as the linear scale for the voltage and current co-ordinates, we must take the square root of the applied power for our scale. In view of this, it might be argued that it would be better practice to refer to the threshold in terms of root-power, rather than power itself.

The conclusions to be drawn must refer back to the two distinct uses to which shock has been put. As a negative reinforcer, where the experimental animals run over a grid, it is obviously of little use to control either the voltage or the current, as this leaves the wattage to vary with the area of contact. With a constant power stimulator the subjective intensity can vary only if skin thickness varies or if the animal gets its feet wet. In these circumstances it would therefore pay to standardize shock as a negative reinforcer in terms of wattage. There is also a strong argument in favour of a.c. rather than d.c. stimulation. Since shock is felt only on the make or break with d.c. stimulation, the degree of negative reinforcement can be minimized by 'freezing' on the grid. With a.c. stimulation it is impossible to escape the shock this way.

In the 'therapy' field the implications could be more serious. The occasional unfortunate demise of a patient who is receiving the 'therapy' suggests that there are times when the stimulus has exceeded its intended effect. In such cases it might be instructive to consider how much power was dissipated, and whether skin resistance was particularly low. There is presumably a limit to the amount of power which may be dissipated in the brain, beyond which the patient fails to survive.

Thanks are due to Mr J. W. Chambers for his skilled maintenance of an unusual piece of circuitry, and to Miss O. Kosviner for her conscientious fortitude as a subject in a most tedious and exacting experiment.

#### REFERENCES

- ADRIAN, E. D. (1928). *The Basis of Sensation*, pp. 63-4. London: Christophers.
- CURTIS, H. J. (1950). Bioelectric measurements. In: F. M. Ueber (ed.), *Biophysical Research Methods*, pp. 233-70. New York: Interscience Publishers.
- FORBES, T. W. & BERNSTEIN, A. L. (1935). The standardization of 60-cycle electric shock for practical use in psychological experimentation. *J. Gen. Psychol.* **12**, 436-42.
- GILMER, B. v. H. (1937). The sensitivity of the fingers to alternating electric currents. *Amer. J. Psychol.* **48**, 444-9.
- GREEN, R. T. (1958). Threshold for electric shock of the laboratory rat. *Animal Behaviour*, **6**, 72-6.
- GREGORY, R. L. (1956). An experimental treatment of vision as an information source and noisy channel. In: C. Cherry (ed.), *Information Theory. Third London Symposium*, pp. 287-99. London: Butterworth.
- HILL, H. E. et al. (1952). Relationship of electrically induced pain to the amperage and the wattage of shock stimuli. *J. Clin. Invest.* **31**, 464-72.
- KATZ, B. (1939). *Electric Excitation of Nerve*. Oxford University Press.
- SOLOMONS, L. M. (1900). A new explanation of Weber's law. *Psychol. Rev.* **7**, 234-40.
- TANNER, W. P. & SWETS, J. A. (1954). A decision-making theory of visual detection. *Psychol. Rev.* **61**, 401-9.

(Manuscript received 25 January 1961; revised 4 August 1961)



## SOME INVESTIGATIONS OF PERCEPTION OF MOVEMENT AND RELATED DEPTH PHENOMENA

By J. L. ZAJAC

*University of Edinburgh, Scotland*

Investigations of perception of real movement have been carried out using a special apparatus designed by the author. Movement of a pendular character was produced by placing a prismatic 'Chance Crookes' glass in front of one eye of an observer looking through it at a vertical rod, and rotating the glass. Amplitude and direction of the movement were a function of the distance of both the glass and the rod from the eye, of the thickness and prismatic angle of the glass, and of the direction and angle of turn of the glass in relation to the eye. When angle of turn or speed of rotation, or both, are gradually increased, other factors remaining constant, observations concerning movement, in both monocular and binocular vision, varied following fixed sequences for monocular and binocular vision. The experiments consisted in noting the kind of observation and recording the related angle of turn and number of rotations of the prismatic glass; from these data mean speeds of the movement of the image of the rod were calculated. Observations concerning depth phenomena accompanying various kinds of perceived movement were also made.

### INTRODUCTION

Many investigations on perception of movement have been carried out, but perception of real movement did not at first arouse great interest. The earliest investigations dealt with 'apparent' movements or 'illusions', i.e. with phenomena in which movement was perceived where objectively no movement took place, or where the movement was of quite a different character from that perceived.

At the beginning of the nineteenth century interest was aroused in the Purkinje phenomenon which was not only a visual phenomenon but was related to changes in the semicircular canals of the ear. Then came stroboscopic and cinematographic observations. These dealt with perceived movements which were related to changeable static elements quite different in character from those actually perceived. Some of these movements were accompanied by complicated phenomena of 'flicker'.

In modern psychology it was Wertheimer (1912) who introduced a new type of investigation of apparent movement. In place of complicated stroboscopic or cinematographic displacements he studied the appearance and disappearance of simple figures (straight or curved lines) using the tachistoscope.

Details of apparatus and methods used in the study of perception of movement have been summarized by several authors (e.g. Neff, 1936; Boring, 1942; Graham, 1951; Gibson, 1954; Schober, 1954).

The present investigations are concerned with real, although not objective, movement, and the apparatus and method used were different from those used in other investigations; both were arrived at by chance.

In the course of experiments on depth perception with glass plates and colour filters the question arose of what changes would occur when smoked glasses were used in place of ordinary glass plates. One school of thought claims that depth perception depends on the illumination and that various colour filters yield different depth perception only because different values of illumination for each colour bring



different localization in the third dimension. This theory is supported by the 'Pulfrich effect' (Pulfrich, 1922\*).

In order to investigate the relation between the degree of illumination (or brightness) and depth perception, the 'Chance Crookes' Alpha, A<sub>2</sub>, B, B<sub>2</sub> glasses were acquired, but at first consistent results could not be obtained. However, some other interesting observations were made. It is known that if an ordinary glass plate or colour filter is placed in front of the right eye, in binocular vision, turning the glass plate away from the nose shifts the stereoscopic image to the left and nearer to the observer, and vice versa when the plate is turned towards the nose. When some of these 'Chance Crookes' glasses were used the observations were not consistent.

It was found, however, that when the glasses were turned through a large angle (50°–70°) and then rotated around their centre in their plane, the stereoscopic image of the part of the fixation rod covered by the glass moved to the left and then to the right, or vice versa, and at the same time towards the observer and then away from him. The mystery of these phenomena was solved when the thicknesses of different parts of the glasses were measured. It was found that the surfaces of those glasses were not quite parallel, different parts differing in thickness although the surfaces were plane; these particular glasses thus produced a prismatic effect, the refracting angles being very small. This feature of these particular glasses was of course not normal: their prismatic character was obtained by some mistake during their elaboration in the workshop. Table I shows the various features of these glasses.

In experiments carried out with these particular 'Chance Crookes' glasses, one obtained very interesting results. It was found in preliminary experiments, as was to be expected, that by turning these plates through a greater angle one obtained much greater lateral shifts than by turning much thicker glass plates with parallel surfaces. It has also been shown that the greater the distances of the plates from the object, the greater also were the changes in direction of vision and of perceived depth. In monocular vision lateral shifts were quite distinct and consistent, while changes in depth were not so distinct and persistent as in binocular vision. For example, with binocular vision with a glass with a difference of thickness of 0.29 mm. and with a thickness in the middle of 1.77 mm. one obtained a difference of distance of 24 in. or greater, depending on the angle of turn of the glass, from the object viewed. This gave the relation between the difference of thickness and the difference in distance as more than 1:2,000. The following factors were shown to be involved: (1) geometrical distance of the object from the observer; (2) distance of the object from the glass; (3) distance of the glass from the observer; (4) thickness of the glass at the point of passage of the visual line connecting the centre of the eye and the object; (5) the prismatic angle, i.e. difference of thickness of the glass; (6) the angle and direction of turn of the glass in relation to the eye; and (7) direction and speed of rotation of the glass. It is to be assumed that all these factors must be related by some formula which could be found theoretically, as well as by experimental procedure.

\* The Pulfrich effect can be described as follows: a pendulum-bob oscillating in a frontal plane will be seen with both eyes as moving in an ellipse whose plane is perpendicular to the frontal plane when a smoked glass or a coloured filter is placed in front of one eye, or if smoked glasses transmitting sufficiently different amounts of light, or two filters of different colours, are placed in front of the two eyes.

In connexion with the above observations several experiments were carried out with the aid of apparatus devised by the author and constructed by Captain W. Bogucki.

#### APPARATUS

The apparatus is shown in Fig. 1. The main apparatus consists of a circular base (*g*) in the centre of which a vertical axis (*v*) was built and a small motor (*m*) the revolutions of which are transmitted by a nylon thread (*n*) to an upper horizontal axis (*h*), whose end is placed vertically over the vertical axis (*v*). On this end one places a cylinder with a glass plate, a colour filter or a 'Chance Crookes' glass (*f*). This glass can be moved in three directions: (1) towards and away from the observer; (2) around the vertical axis (*v*) of the apparatus, i.e. turned towards or away from the nose of the observer; and (3) rotated around the horizontal axis (*h*) on which it is placed. The movements (1), (2) and (3) can be made by the hand of the observer or by an operator; the rotation movement (3) also by electric current by means of the rheostat (*d*).

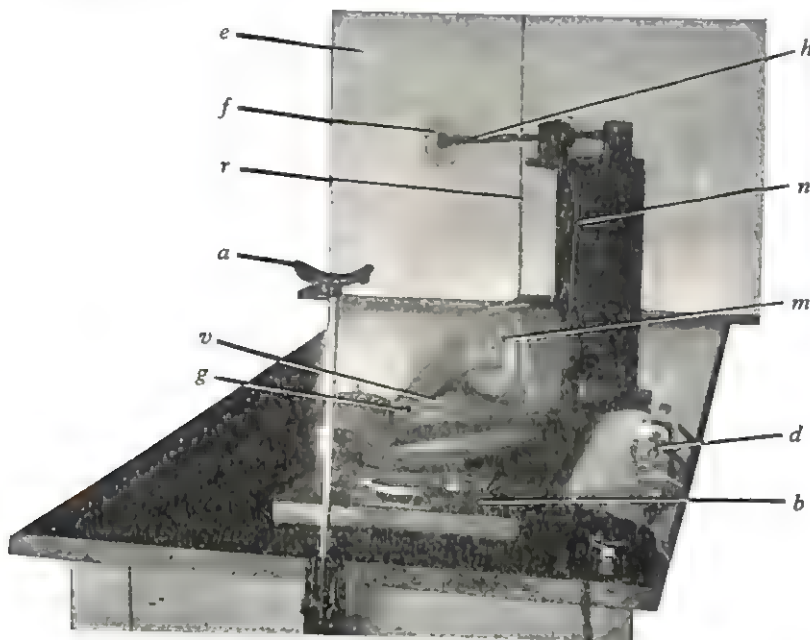


Fig. 1

The rheostat enables one to change the frequency of revolutions from about 20 r.p.m. to over 1,200 r.p.m. Pending the incorporation of a revolution counter the number of rotations of the glass can be counted by measuring the speed of rotation of the transmission-thread with a stop-watch. The motor and the axes are mounted on a base (*b*) about 4 ft. long, in such a way as to be movable backwards and forwards, and different objects, i.e. rods (*r*), may be mounted on it. A head-rest (*a*) and a back screen (*e*), limiting the field of vision, complete the apparatus.

This apparatus enables one to carry out investigations of visual perception of movement, both monocularly and binocularly, to study depth perception in relation to that movement, to detect and eventually to measure differences in thickness of glass plate (differences of prismatic character), using stereoscopic methods, and to study 'flicker' phenomena and the Pulfrich effect.

When we rotate about the horizontal axis a prismatic glass plate placed, for example, in front of the right eye, we can distinguish its different positions, four

Different phenomena are observed in binocular vision with a prismatic glass placed in front of one eye. When the angle of turn of the glass or the speed of rotation is increased while other factors are held constant, the phenomena are observed in the following sequence: (1) The image of the rod moves obliquely to the median plane. Amplitude and direction of this movement are unequivocally defined. The single image of the moving rod is the result of the stereoscopic fusion of two images, one, stationary, belonging to the eye without a glass in front of it and the other moving as a result of the rotation of the glass plate. (2) One image of the rod begins to be separated from the other; the image of the eye without a filter follows the movement of the other for some time, but then stops. (3) Two separate images are seen, one moving and the other stationary. (4) Three images are seen, one belonging to the eye without a glass in front of it, relatively steady and stationary, and two others similar to those mentioned above at points (2) to (6) inclusive, for monocular vision; and, in binocular vision, in the same sequence, but with the difference that these two images are rather hazy in comparison with those seen with one eye.

Instead of a vertical rod we can use one with its top inclined towards the observer. The main phenomena are similar to those described above, but with one eye the inclination of the top towards the observer is much less pronounced than in the case of binocular vision; sometimes it appears vertical but the upper part seems to be thicker than the lower part. The character of the movement in monocular vision may be changing or undefined, but when we see, through the glass, the movement of the rod's base, which goes up and down and moves sideways, then the perceived character of the movement of the image of the rod becomes definitely established: the rod is seen moving along a circle or ellipse and recedes from the observer when the base is going up, and approaches him when it goes down. The rotating movement is anticlockwise when the glass is turned towards the observer's nose, and clockwise when the glass is turned away from the nose.

In binocular vision one can in this case observe phenomena in the following sequence: (1) One rod moving obliquely, but more inclined with its top towards the observer than in the case of monocular vision. (2) The rod is seen splitting into two images. (3) Two images are seen intersecting at a point which is fixated. Both are still more inclined than each seen separately by the corresponding eye. The moving image seen through a glass is, so to speak, dragging the other one behind it for a part of the way. (4) The dragging effect ceases, but the greater inclination of both images or only of the stationary one, which is rather surprising, persists. (5) The images are seen in the position nearing the vertical. (6) Three images appear.

Table 3 illustrates some of these phenomena. It contains results of one series of experiments, when the distance of the glass plate from the right eye of the observer was 10 in. and a vertical grey rod, 6 mm. in diameter, was placed 30 in. behind the glass plate in the vertical position. The glass was turned with its left edge towards the observer's nose. The phenomena to be noted by the observer in monocular vision were: (a) splitting of the image into two, (b) two images moving, and (c) two images flickering. In binocular vision: (a) one image stationary, the other moving obliquely to the line of sight backwards and forwards; (b) the other image moving from side to side in the frontal plane; (c) one image stationary, the other two moving; and (d) one image stationary, the other two flickering. In this experiment the glass B was turned



through angles of 55°, 60°, 65° or 67·5° and by changing the speed of rotation one tried to make the observer perceive one of the phenomena enumerated above. One measured the corresponding speed of rotation and in this way one obtained data which allowed the computation of the numbers of rotations per minute and the mean speeds of the movement. The results are contained in Table 3. Three persons, OD, OE and OF, served as observers and for each situation four or five measurements were made, two of ascending, two of descending series, and eventually one to control the uncertain observations.

Table 3. Observations obtained with 'Chance Crookes' glass B

Observations	Angle of turn of the glass (deg.)	OD		OE		OF	
		r.p.m.	Speed (cm./sec.)	(r.p.m.)	Speed (cm./sec.)	r.p.m.	Speed (cm./sec.)
		Monocular vision					
Image splitting into two images	55	396	14.5	375	13.75	396	14.5
	60	336	13.5	356	16.6	316	14.7
	65	274	17.4	285	17.1	303	14.7
	67.5	269	18.8	259	18	230	16.1
Two images moving	55	462	17	418	15.3	450	16.5
	60	375	18.5	365	17.1	376	17.5
	65	336	20.2	325	19.4	356	21.3
	67.5	291	20.4	292	20.4	325	21.7
Two images flickering	55	475	17.4	491	18	491	18
	60	462	21.5	418	19.5	466	21
	65	375	23	365	21.6	412	24.6
	67.5	316	21	356	25	356	25
Binocular vision							
One image stationary, the other moving obliquely	55	274	10	223	8.14	348	12.76
	60	172	8	204	9.5	332	15.5
	65	176	10.6	259	15.5	221	13.3
	67.5	159	11	246	17.2	221	15.45
One image stationary, the other moving from side to side	55	348	12.76	317	11.6	385	14.1
	60	316	14.7	285	13.3	348	16.2
	65	274	16.4	285	17.1	279	16.7
	67.5	210	14.7	274	19.1	246	17.2
One image stationary, two others moving	55	406	14.9	385	14.1	407	14.9
	60	365	17	348	16.2	367	17.1
	65	297	17.8	325	19.5	292	17.5
	67.5	285	20	309	21.6	285	20
One image stationary, two others flickering	55	460	16.8	445	16.3	491	18
	60	475	22.2	375	17.5	455	21.2
	65	356	21.3	367	22	375	22.5
	67.5	297	20.8	348	24.3	356	24.9
Distance of the middle of the glass from observer 10 in. dist.							

Distance of the middle of the glass from observer, 10 in., distance of the rod from the middle of the glass, 30 in.

The B glass was used in front of the right eye turned with its left edge towards the nose of the observer.

## DISCUSSION

Table 3 gives the mean number of rotations of the glass per minute and the mean speed in cm./sec. corresponding to particular phenomena for different angles of turn of the 'Chance Crookes' glass B. It confirms the sequence of phenomena mentioned

above resulting from the increase of the speed of movement which is a function of the angle of turn of the glass plate (and with it of the amplitude of movement) and of the number of rotations. We can infer from the data presented in this table that each phenomenon can be observed within certain limits of speed of movement, which on the other hand overlap on different observed phenomena. This overlapping may be attributed to various causes which could not be controlled during this series of experiments, and which therefore will not be discussed here.

In monocular vision with lower speeds than with those corresponding to the splitting of the image into two, we see only one image moving. With greater speeds than for two images flickering, i.e. when CFF (critical flicker frequency) is surpassed, one can observe two stationary images.

In binocular vision with lower speeds than those required to produce one stationary image and the other moving obliquely backwards and forwards, only one stereoscopic image moving obliquely is seen. With greater speeds of movement than those for perceiving one image stationary and two others flickering, one sees three stationary images.

If an object is moving in front of our eyes in pendular motion the speed of this movement must not exceed certain limits, if it is to be seen as an object moving continuously. The limits at which the moving object ceases to be seen as one and splits into two images are different under different conditions of movement and of observation. They depend on many factors, such as illumination, colour, shape of the object, its size, the part of the retina involved, the way of looking at the moving object, e.g. by fixating a stationary point, or by moving the eye in pursuit or reverse movement, etc. Each of these factors influences the limits of speed mentioned above, and therefore it is not surprising that different investigators have obtained different results concerning the limits of speeds. In the present investigations, illustrated in Table 3, the mean speed at which a continuous movement of a vertical rod was still seen was 15.32 cm./sec., or an angular speed of  $8^{\circ} 36'$  per sec., while the mean speed at which splitting of one image into two occurred was 16 cm./sec., or an angular speed of  $9^{\circ}$  per sec., but it must be noted that in some cases even a speed of 19 cm./sec., or an angular speed of  $10^{\circ} 55' 15''$  per sec., gave one image only, and the speed of 13.75 cm./sec., or an angular speed of  $7^{\circ} 44'$  per sec., already gave two moving images. For the present no opinion is offered as to whether these results are due to faults in observation or to mistakes on the part of the operator, or whether the number of observations was insufficient to eliminate errors.

It is noteworthy that the writer's investigations of real movement resulted in the perception of phenomena similar to those of Wertheimer and others who dealt with so-called 'apparent' movement.

We can observe some of the phenomena described above in a natural setting using a rod, stick or similar object held in the hand and moved in pendular motion to right and left. We can then easily discriminate the moment when the one moving object splits into two images as the speed of movement is increased, and then the moment when the two images are seen moving, and even when they are flickering; but it is impossible in this way to obtain two stationary images because we cannot move the hand with sufficient speed to surpass CFF. Another interesting observation can be made when, for example, the tip of a ball-pen is of shining silvery colour

and the rest of it normally coloured red or blue or any other colour; when this pen is moved by hand we can see a continuous stripe made by the moving shining tip, while the rest of the pen is already split into two images with only a blur appearing between them. One must note here that the tip of the pen is moving faster than the remainder of the pen, which would suggest the reverse phenomenon. Similar observations can be made with a white rod with a black ball at its tip and with several stripes painted on it. When the rod is moved with sufficient speed we can observe two images of it moving or flickering, while the image of the dark ball and of the stripes are seen as continuous strips or ribbons. It might be interesting to organize systematic investigations with similar differently coloured objects.

#### DEPTH PHENOMENA CONNECTED WITH MOVEMENT

Observations in monocular vision have established that when only one image of a rod is seen moving, the sequence of observations as the speed or angle of turn is increased might be as follows: first one sees a rod moving in a circle, then in a narrowing ellipse and finally moving from side to side in the frontal plane, or first in an oblique movement changing into a frontal one. In some cases the movement might be seen as clockwise, anticlockwise, or alternating; the direction of movement is not an unequivocal one and may be changing. But if the observer sees the base of the rod moving up and down, then the direction of movement is definitely established and unchanging: the image comes nearer when the base goes down, and recedes from the observer when the base goes up.

Nothing of this kind is observed in binocular vision when one stereoscopic image is moving obliquely to the line of sight: in this case the position of the base does not matter at all: some observers see the image making a very narrow ellipse, which can correspond to the Pulfrich effect. From these observations one could infer that here the so-called 'height' clue in described conditions could define the character of depth perception only in monocular but not in binocular vision. Perhaps outside the practical range of convergence the 'height' clue might be important for depth perception in binocular vision in the same way as in monocular vision.

In binocular vision with increased speed of rotation or increased angle of turn the single stereoscopic image splits into two double images: one, belonging to the eye which looks through the glass, moves obliquely as the stereoscopic image did previously, the other, seen by the other eye, moves first along part of the way as if dragged by the other image and then stops. This could be defined as an 'induced' effect; possibly eye movements play a role here. It is important and interesting to note this oblique movement of the double image seen by the eye with a glass in front of it, because this movement is similar to that of the stereoscopic image perceived with smaller speeds of rotation or smaller angles of turn. This is proof that within certain limits depth perception of double images is subject to the laws of stereoscopic depth, although here the disparity of both images greatly surpasses the so-called Panum areas. This has been shown also in the work of Ogle (1952) and of the writer (Zajac, 1956).

On the other hand, the movements of one image influence the behaviour of the other image so that the other image follows the first one for some time. It is a particular



case of an influence exerted by occurrences in one eye on perceptions by the other eye. Other phenomena connected with this influence have been discussed by the writer in other papers (Zajac, 1956, 1959).

Interesting phenomena occur when the rod is placed about 30 in. behind the glass and its top inclined towards the observer at about  $45^\circ$ . In monocular vision this inclination is perceived rather indistinctly both with and without a glass plate in front of the eye. It is perceived rather as vertical or near vertical but of uneven thickness, thicker at the top and thinner at the bottom, both when it is stationary and when it is moving. Quite different phenomena occur in binocular vision. The inclination is very well perceived, not only when a single stereoscopic image is seen but also when it is splitting into two images; even the stationary image seen by the eye not covered by the glass appears for some time as inclined at the top towards the observer and—what is more curious—it is still seen as inclined when the moving image belonging to the eye covered by the glass is already nearing the vertical. With still greater speed of rotation or greater angle of turn, two or three images are seen as vertical and not inclined.

The experiments described cast some light on the value for depth perception of certain clues (or empirical factors, as some authors call them). Three of these factors will be discussed: size, overlay, height.

In the situations where in binocular vision we see only one stereoscopic image moving obliquely to the median plane, we observe the image increasing in size when it is moving away from the observer and diminishing in size when it is coming nearer. We can infer from this that the influence of size on depth perception is not very important in comparison with the factor of disparity, and is sometimes negligible in stereoscopic vision in the near visual field. As regards monocular vision, the present investigations afforded no evidence as to whether slight changes in size could influence depth perception of moving monocular images.

The second factor, overlay or overlapping, has been shown in our experiments on movement to play no role in depth perception where stereoscopic images are involved. This can easily be proved by placing a rod near the white screen and observing it binocularly through the prismatic glass in front of one eye turned through a large angle ( $65^\circ$ – $75^\circ$ ). When this glass was slowly rotated, it appeared as if the stereoscopic image of the rod were moving across the surface of the background and going behind the part of the background not covered by the glass. The shadow of the rod thrown on this background or a line drawn on it and seen outside the glass, together with the whole background, were perceived nearer than the stereoscopic image. With monocular vision this impression was never obtained; the image of the rod was seen always in front of the background, but was perceived much nearer to it than geometrical distance would suggest. This latter effect is an aspect of the factor of overlapping which is not generally mentioned.

The third factor, height, does not influence depth perception in binocular vision in the near field of view in our experiments, as we have seen above, and cannot compete with the factor of disparity, but is of some importance and sometimes decisive in monocular vision. In our case it determined the direction of perceived movement and corresponding depth phenomena.

All these remarks concern the near field of vision where convergence, and with

it disparity, play their role in binocular depth perception. Outside the limits of practical convergence (100–300 m.) binocular vision, as regards depth perception, resembles monocular vision and then the roles of the various empirical factors are similar for both.

One final remark about the relief phenomena perceived in binocular vision in the case where one sees one stereoscopic image moving obliquely: when this image is moving away from the observer it grows and also changes its shape; it assumes a somewhat elongated form of an almost slanting surface, or of a 'wall' instead of a round rod. This is undoubtedly a result of the fact that the two images, one seen through a glass and another without it, change their relative size when the glass is rotated. Different parts of the glass are of different thickness and with changing thickness the retinal images also change their size. The size of the image belonging to the eye without a glass changes only slightly, owing to changes of accommodation of the eye, when changes in convergence occur due to the rotation of the glass in front of the other eye.

The results of these investigations confirm several of the principal laws of depth perception: (1) Binocular depth perception within the limits of practical convergence is different from monocular depth perception. (2) The so-called empirical factors, such as size, overlay, height, do not play any important role in stereoscopic depth perception in the near field of vision, or at least their importance is small in comparison with that of correspondence and disparity near the horopter. (3) The images in one eye influence, in some cases, depth perception of the images in the other eye.

I wish to thank Professor James Drever for giving me the opportunity of carrying out these investigations; the Senatus Academicus of the University of Edinburgh, the Executive Committee of the Carnegie Trust for the Universities of Scotland and the Committee of the Earl of Moray Endowment Fund for grants in aid which made this study possible; Captain W. Bogucki who constructed the apparatus used in the experiments and many students who served as observers.

I am especially indebted to the late Miss Czesława Ramult, M.A., for her assistance in this work, and for acting as operator when I served as an observer.

I also thank Dr Catherine J. D. Jarvis of the National Research Council of Canada for revising the text of this paper.

## REFERENCES

- BORING, E. G. (1942). *Sensation and Perception in the History of Experimental Psychology*, pp. 588–99. New York: Appleton-Century-Crofts.
- GIBSON, J. J. (1954). The visual perception of objective motion and subjective movement. *Psychol. Rev.* **61**, 304–13.
- GRAHAM, C. H. (1951). Visual perception. In S. S. Stevens, *Handbook of Experimental Psychology*, pp. 895ff. New York: Wiley.
- NEFF, W. S. (1936). A critical investigation of the visual apprehension of movement. *Amer. J. Psychol.* **48**, 1–42.
- OGLE, K. N. (1952). On the limits of stereoscopic vision. *J. Exp. Psychol.* **44**, 253–9.
- PULFRICH, C. (1922). Die Stereoscopie im Dienste der isochromen und heterochromen Photometrie. *Naturwissenschaften*, **10**, 553–64, 569–74, 596–601, 714–22, 735–43, 751–61.
- SCHOBER, H. (1954). *Das Sehen*, 2, pp. 124–8, 374–9. Leipzig: Fachbuchverlag.

- WERTHEIMER, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.* **61**, 161-265.
- ZAJAC, J. L. (1956). Depth perception of double images in the vicinity of other images. *Acta Psychol.* **12**, 111-29.
- ZAJAC, J. L. (1959). Depth perception of stereoscopic images resulting from fusion of crossed and uncrossed double images. *Amer. J. Psychol.* **72**, 163-83.

*(Manuscript received 14 December 1960)*



## CONTROL, DEFENCE AND CENTRATION EFFECT: A STUDY OF SCANNING BEHAVIOUR\*

By RILEY W. GARDNER AND ROBERT I. LONG

*The Menninger Foundation*

Studies of individual consistencies in cognitive behaviours by Gardner and others led to postulation of a cognitive control principle called extensiveness of scanning that could in part determine the apparent size of standard stimuli in a relatively difficult size estimation test. In the present study, electro-oculography was employed to obtain precise measures of the scanning strategies 60 subjects employed in four size estimation tests. In addition to providing evidence of consistent individual scanning syndromes, the study showed that extensiveness of scanning is a determinant of apparent size in the relatively difficult size estimation test in question. The study also confirmed Piaget's hypothesis (see Piaget, 1961) that the apparent size of a stimulus is a function of the duration of a single centration upon it. Subjects were shown to be consistently different from each other in these centration effects, independent of the significant negative correlation of the effect with age. These individual differences were significantly associated with the apparent size of standard stimuli in a relatively simple size estimation test. A predicted relationship between extensiveness of scanning and the number of conceptually distant responses produced in a free association test was not confirmed. Predicted relationships were confirmed, however, between extensiveness of scanning and the strength of the defence mechanisms of isolation and projection and the degree of generalized delay, judged on the basis of Rorschach test protocols. Predictions of specific aspects of response to the Rorschach inkblots were also confirmed.

### I. INTRODUCTION

In reporting an exploratory study of individual consistencies in cognitive behaviour conceived in terms of dimensional principles of cognitive control, Gardner *et al.* (1959) described a factor for men that appeared to represent the extensiveness with which persons scan both external and internal informational arrays. Subjects assumed to be extensive scanners showed minimal overestimation, or underestimation, of standard stimuli in one type of size estimation test. They also seemed to show relatively great recall of incidental material in a colour-word test and relatively extensive sampling of internal fields of ideas, as exemplified by production of many free associations distant in meaning from two stimulus words. In that study and in a subsequent study (Gardner, 1961) of 80 women subjects, constant error in the type of size estimation test employed was assumed to be determined in part by the manner in which the subject scanned the standard and comparison stimuli. These reports drew upon Piaget's observations (Piaget, 1947, 1961; Piaget, Vinh-Bang & Matalon, 1958; Piaget *et al.*, 1942-56, 1959) concerning relations between 'errors of the standard' (overestimations of constant stimuli) and the patterning of attention and described new results suggesting that individuals are consistently different from one another in errors of the standard under free scanning conditions. Neither study, however, included actual measures of subjects' scanning behaviours. Recent developments in

\* This investigation was supported by research grant M-2454 from the National Institute of Mental Health, United States Public Health Service. We are indebted to Mr Walter Kintsch, who helped us to develop the procedures and scores used to assess scanning behaviours and contributed to collection of the data, and to Mr Leander Lohrenz and Mr Robert Schoen, who evaluated the Rorschach Test protocols. We are also indebted to Dr Gardner Murphy and Prof. Jean Piaget for thoughtful comments on the work described in this report.

electro-oculography (see for example, Ford & Leonard, 1958; Ford, White & Lichtenstein, 1959; Kris, 1960; Shackel, 1960; Shackel & Davis, 1960) made it possible to obtain precise measures of the gross horizontal eye movements involved in the size judgements employed in the present study.

According to Piaget, the magnitude of a stimulus in the centre of the attentional field is automatically overestimated. In addition to being a function of the magnitude of the stimulus, the amount of overestimation is a function of at least five other determinants (Piaget, 1961, pp. 113-14), including the duration of centrations upon it. In laboratory experiments, errors of the standard are assumed to occur because of the accrual of relatively great centration times upon constant stimuli. In the course of development, persons apparently evolve strategies of attention deployment (decentration) that serve to counteract these centration effects. In adults, extensiveness of scanning could therefore lead to minimal overestimation, or underestimation, of standard stimuli in size estimation tests (see Gardner *et al.* 1959; Gardner, 1961).

Gardner & Long (1960*a, b*) confirmed observations of Piaget and his associates by showing that errors of the standard of considerable magnitude occur in judgements of an inverted-T figure and a reversed-L figure. In the present study, inclusion of a size estimation test in which all subjects were required to judge the size of one standard stimulus after single centrations of prescribed durations provided a test of Piaget's hypothesis that the apparent size of a stimulus is a function of the duration of centration upon it. Inclusion of this test also made it possible to evaluate the authors' hypothesis that individuals differ consistently in the amount of centration effect *per se*, independent of scanning. Confirmation of this hypothesis would make it possible to assess the *relative* contributions of scanning strategies and centration effects to the apparent size of standard stimuli in free scanning tests.

In the study by Gardner *et al.* (1959), the apparent link between extensiveness of scanning and the strength and predominance of isolation (Freud, 1926) in the pattern of defensive structures was of greater theoretical importance than the specific test correlates of size estimation responses. According to psychoanalytic theory, isolation—the splitting of affect from idea and idea from idea—is the defence mechanism involved in 'obsessive-compulsive' behaviours. In the earlier study referred to, a significant proportion of subjects who were judged to employ isolation as their predominant defence (i.e. 'obsessive' subjects) were among those thought to be extensive scanners. Other extensive scanners, however, did not rely primarily on this defence. Confirmation of these findings could indicate a theoretically important link between cognitive structures serving purposes of defence and structures called cognitive control principles, which seem apparent in response to various relatively 'neutral' adaptive requirements. If one assumes that the cognitive controls characterizing an individual are general principles of cognitive organization, extensive scanning, with its implied emphasis upon caution, exactness, and delay, could be linked to more than one defence. Although presumed to be rooted in conflicts essentially different from those leading to isolation, the defence of projection (see Freud, 1922) is often inferred in part from a form of reality contact manifested by cautious, hyperalert exactness in the individual's approach to both persons and things. Relationships between extensiveness of scanning and both isolation and projection were assessed in the present study on the basis of responses to the Rorschach Test.

Consistent individual differences in extensiveness of scanning, if demonstrated, would imply different degrees of *delay* between confrontation with an adaptive problem and commitment to a response. In a size estimation test, for example, extensive scanning of standard and comparison stimuli before committing oneself to a judgement should extend decision time. If degree of delay is a generalized characteristic of the individual's cognitive functioning, it should also be apparent in response to relatively 'unstructured' situations. Responses to the Rorschach Test were also used to test this general assumption.

## II. HYPOTHESES

Because of these theoretical and empirical considerations, the present study was designed to test the following specific hypotheses: (1) Individuals differ consistently in scanning strategies, as indicated by the patterning of eye-movements in size estimation tests. (2) The apparent size of a standard stimulus is a function of the duration of a single centration upon it. (3) Individuals differ consistently in centration effect *per se* (i.e. in the experienced size of a standard stimulus under identical viewing conditions). (4) Repeated judgement of a single standard stimulus leads to a progressive increase in its apparent size. (5) Extensiveness of scanning [e.g. in the form of number of centrations on standard (constant) stimuli] is negatively related to the apparent size of standard stimuli under free scanning conditions. (6) Centration effect is positively related to apparent size under free scanning conditions. (7) Extensiveness of scanning in size estimation tests is negatively related to error of the standard in judgements of inverted-T figures. (8) Extensive scanners offer greater percentages than limited scanners of conceptually 'distant' responses in a free association test. (9) Extensive scanners rely more than limited scanners upon the defence of isolation. (10) Extensive scanners rely more than limited scanners upon the defence of projection. (11) Extensive scanners show significantly more generalized delay than limited scanners in a relatively unstructured test. (12) Extensive scanners spend more time than limited scanners examining the Rorschach inkblots before offering first responses. (13) Extensive scanners are more likely than limited scanners to indicate concern with the symmetry of the Rorschach inkblots.

## III. METHOD

### (i) Subjects

The sample of 60 women included housewives, secretaries, and business college students ranging in age from 17 to 44, mean age 22 years 11 months.

### (ii) Size estimation tests

Four size estimation tests were administered to all 60 subjects. Three of these tests investigated response variations attributable to the nature of standard and comparison stimuli, the size of standard stimuli, and the number and nature of judgements under free scanning conditions. The fourth test provided a measure of centration effect in single centrations of four prescribed durations upon a standard stimulus. The four tests were given to each of the 60 subjects in the order in which they are described here. The standard (constant) stimuli were always on the subject's left. In each test, eye-movement recordings were obtained with an oscillographic inkwriter by means of an Offner Type P Dynograph Assembly equipped for curvilinear recording. Chlorided pure silver electrodes embedded in plastic cups filled with salt paste were held at the level of the pupils



on the right and left temples by means of metal clips with plastic contacts that were held in place by an adjustable headband. The experimenter registered the opening and closing of the subject's eyes at the beginning and end of each judgement and all stops in the subject's adjustments of comparison stimuli by pressing a telegraph key that activated a marker pen.

(a) *Free scanning tests*

*Test I.* This test was identical with that used by Gardner (1961). It was identical with size estimation Test I described by Gardner *et al.* (1959), except that three 40 mm. disks, rather than disks of varied sizes, were judged. The following disks were used: grey, 7 g.; black, 7 g.; heavy grey, 67 g. The subject made four judgements of each disk from starting-points of 30, 50, 48, and 32 mm. by turning a crank that varied the size of a light circle appearing on a glass screen to the right of the standard stimuli (a point source of light was used). Before the test, the experimenter attached the electrodes and adjusted the subject's head in a head-rest so that her vision was centred halfway between the standard and the comparison stimuli, which were 7 in. apart. Viewing distance was 14 in. A square object was used to show the subject how to hold the standard disks at the same height and in the same plane as the variable circle of light. She was instructed to keep her eyes closed before each judgement and to close her eyes as soon as she was satisfied with each judgement.

Because of the several differences between the standard stimuli and the adjustable light circle (primarily the weight, volume, colour, and relatively diffuse contour of the disks when held in the hand), this test is both more complex and more difficult than the two free scanning tests described below. The fact that these disks have been overestimated by most subjects in previous studies may be attributable not only to scanning patterns and/or centration effects but also to the illusion of enhanced size produced by the disk properties listed above.

*Test II.* The Test I apparatus, with an incandescent rather than a point source of light, was used to produce an adjustable light circle. An identical apparatus was used to produce a 40 mm. standard light circle. The positions of the stimuli and viewing distance were the same as in Test I. The subject made four judgements from the Test I starting-points.

*Test III.* This test was like Test II, except that the standard circle of light was 30 mm. in diameter and the subject made ten consecutive judgements from the same 22 mm. starting-point.

(b) *Controlled centration test*

*Test IV.* In this test, the standard light circle was 50 mm. in diameter. Each subject was allowed a single centration of predetermined duration upon the standard before each adjustment of the comparison circle from the 42 mm. starting-point. She viewed the standard stimulus for 1, 2, 3, and 4 sec. in one of four orders: 1234 ( $N = 17$ ), 2143 ( $N = 14$ ), 3412 ( $N = 16$ ), 4321 ( $N = 13$ ). Inter-trial interval was one minute. A practice session prior to the test insured that the subject could follow the instruction to open her eyes and look at the area in which the standard stimulus would appear immediately after the experimenter tapped with a pencil and the instruction to look at and adjust the comparison light circle immediately after a second pencil tap.

(c) *Scanning scores*

For each subject, mean scanning scores per judgement in Tests I, II and III were derived from the following basic scores: (1) *judgement time*: the number of seconds that elapsed between the subject's opening her eyes and completion of her last adjustment of the comparison stimulus. (2) *Number of stops*: the number of times the subject stopped cranking in the course of adjusting the comparison circle. (3) *Number of centrations on standard*: the number of centrations upon the standard during the judgement time. (4) *Time on standard*: the sum of centration times on the judgement time the subject spent in centrations on the standard. (5) *Percentage of time on standard*: the percentage of the mean number of seconds per centration on the standard. (6) *Time per centration on standard*: the mean number of seconds per centration on the standard during the judgement time. (7) *Redundant scanning time*: the number of seconds the subject spent 'checking' her judgement after the judgement time, before closing her eyes to indicate she had completed her judgement.

(ii) *Error scores*

Mean constant error (error with signs) per judgement was computed in millimetres for each subject for each size estimation test. For correlational purposes, use of a single constant error score for Test I seemed justified by the fact that Pearson  $r$ 's between constant errors with the three different stimuli employed ranged from 0.79 to 0.88. An analysis of variance indicated that order effects were not significant in Test IV, so that mean constant error scores for the four groups of subjects could be pooled.

(iii) *Inverted-T test*

The inverted-T procedure used to measure error of the standard was that employed by Gardner (1961) and Gardner & Long (1960a). In one figure the horizontal line, in the other figure the vertical line, served as the standard (constant) stimulus. Scores from this test were not used, however, since inquiry revealed that a number of the subjects became familiar with this form of the vertical-horizontal illusion in a course in dress design and were consciously adjusting their settings of the comparison stimuli to compensate for the illusion.

(iv) *Additional tests administered later to extreme groups*

Following preliminary analysis of the data, groups of ten extensive and ten limited scanners were selected on the basis of mean rank in judgement time, number of stops, and redundant scanning time in size estimation Test I and asked to return for further testing. The interval between original and further testing ranged from 25 to 97 days, the mean being 60 days. The four tests described below were administered.

*Repetition of Size Estimation Test I.* An exact repetition of this test was performed.

*Size Estimation Test V.* This test was identical with Test I except that subjects judged the size of a black wooden disk, 40 mm. in diameter, affixed to a flesh-coloured board attached to the apparatus producing the adjustable circle of light. The disk was in the same position as the hand-held disks of Test I. This test was designed to control for effects the holding of Test I disks might have had on scanning and apparent size. The scanning and error scores were the same as those derived originally from Tests I-III.

*Reversed-L Test.* Two reversed-L figures, unfamiliar to the subjects, were used to assess errors of the standard in judging horizontal and vertical lines. The figures were strips of white plastic tape,  $\frac{1}{8}$  in. wide, affixed to a 38 by 44 in. black cardboard background. In one figure the horizontal line, in the other the vertical line, served as the 100 mm. standard. Viewing distance was 10 ft. The room was completely dark except for a diffuse circle of light illuminating the figure. Only the white plastic strips were visible to the subject. The subject's eyes were open only during adjustments of the comparison stimulus. She was asked to adjust it until it looked equal in length to the standard. Ten subjects judged the horizontal standard first, ten the vertical first. Following dark adaptation, the subject made four judgements (ADDA order) of the standard line in one of the figures by turning a control knob that activated a selsyn on the back of the figure. This apparatus controlled movement of a covering black cardboard strip which changed the visible length of the comparison stimulus. Ascending and descending starting-points were equal distances from the anticipated mean adjustment based on performances of pilot subjects.

The 10-min. interval between judgements of the two figures was filled by a kinesthetic after-effect test, during which the subject was blindfolded. The subject then made four adjustments of the standard line in the other figure from ascending and descending starting-points equal distances from the anticipated mean adjustment.

Since the error of the standard decreases the negative constant error with the horizontal standard and increases the positive constant error with vertical standard, constant error, vertical standard, plus constant error, horizontal standard, is the total error of the standard with the two forms (see Gardner, 1961; Gardner & Long, 1960a, b). Order of presentation had a negligible effect on the mean error of the standard for the two groups, so that the score described here could be used without correction.

*Free Association Test.* The subject was required to associate for 3 min. to each of two stimulus words, 'Dry', and 'House'. The score was the percentage of responses to the two words that were classified 'most distant' from the stimulus word according to the seven-category classification scheme used by Gardner *et al.* (1959).

*Rorschach Test.* This test was administered and scored as suggested by Rapaport, Gill & Schafer (1946). The senior author initially evaluated the twenty protocols in terms of the strength of isolation. It was readily apparent that none of the twenty female subjects could be said to rely primarily on isolation. In contrast to the typical records of truly obsessive persons, for example, the protocols were generally brief, limited in number of content categories, and lacking in strained attempts at 'precision' by verbal elaboration of responses. In spite of this sampling limitation, two independent judges were asked to divide the Rorschach protocols of the twenty extreme subjects into 'high' and 'low' groups of ten protocols each according to the degree of reliance on the defence of isolation.

Subsequently, the judges divided the twenty Rorschach protocols into two groups of ten in terms of the strength of the defence of projection. Schafer's (1954) suggestions concerning assessment of projection from Rorschach protocols served as a guide in these ratings.

Each of the judges was also asked to divide the twenty Rorschach protocols into two groups of ten in terms of the degree of generalized delay evident in the responses. 'Delay' was defined in the psychoanalytic sense. For example, a tendency to give easy popular responses indicating a minimum of blot articulation, as well as a tendency to make impulsive comments or exclamations, could contribute to placing a subject at the 'low delay' end of the continuum.

By Fisher's Exact Test (a one-tailed test), the probability that the judges' agreements on each of the three ratings occurred by chance was 0.09. In the case of each of the ratings, only the fourteen subjects the judges agreed upon were used in the evaluation of hypotheses.

Specific scores obtained to test Hypotheses 12 and 13 were the mean time in seconds before the subjects' first responses to the ten inkblots and the presence or absence in each protocol of indications of preoccupation with the symmetry of the inkblots (including comments on the similarity of the two halves of a blot, discrepancies between the halves, 'mirror images', 'reflections', etc.).

#### IV. RESULTS AND PRELIMINARY DISCUSSION

##### (i) *Mean scanning and error scores*

Table 1 contains means and sigmas of the scanning and error scores. When the variations in stimulus conditions are considered, it is obvious that the scanning scores correspond quite closely, whereas the constant error scores vary considerably. The fact that the 40 mm. disks of Test I were overestimated, in contrast to the 40 mm. light patch of Test II (in which other conditions were identical to those of Test I), could be due to the relative diffuseness of contour, lesser background contrast, or weight and volume of the disks as compared to the sharply defined standard light circles of Test II. For purposes of comparison with constant error in Tests III and IV,

Table 1. *Means and sigmas of scanning and error scores of four size estimation tests\**

(*N* = 60.)

Score	Test I		Test II		Test III		Test IV	
	Mean	Sigma	Mean	Sigma	Mean	Sigma	Mean	Sigma
Judgement time	15.60	5.41	17.39	8.81	15.84	6.29	—	—
Number of stops	0.98	1.27	1.22	1.54	0.81	1.18	—	—
Number of centrations on standard	6.10	3.11	8.38	4.06	6.86	3.13	—	—
Time on standard	3.12	1.76	4.79	3.09	3.96	2.03	—	—
Percentage of time on standard	19.35	5.68	25.87	6.60	24.41	7.15	—	—
Time per centration on standard	0.51	0.14	0.55	0.17	0.58	0.19	—	—
Redundant scanning time	1.16	1.53	1.40	1.56	1.36	2.58	—	—
Constant error (all judgements)	2.12	1.42	-0.40	1.05	—	—	—	—
Constant error (ascending judgements)	1.06	1.47	-0.58	1.25	-0.31	0.67	-1.45	2.02

\* All time scores in seconds.



which included only ascending judgements, constant error in Tests I and II is shown for all judgements and for ascending judgements alone. Starting-position effects are clearly apparent in the differences between the two sets of means for Tests I and II.

(ii) *Individual consistencies in scanning and apparent size*

Table 2 provides clear evidence that individuals differ consistently in scanning strategies (Hypothesis 1) in the three free scanning tests.

Subjects show much less across-test consistency in constant error. Test I, in fact, seems to stand alone. In keeping with Hypothesis 6, constant errors in Tests II (40 mm. light circle standard) and III (30 mm. light circle standard) may in part be determined by the amount of centration effect *per se*, as indicated by the significant correlations with constant errors, Test IV (50 mm. light circle standard, controlled viewing of standard). In view of the correlation of  $-0.42$  ( $P < 0.001$ ) between age and constant error, Test IV, correlations with constant error, Test IV, were also computed with age correlations partialled out. The correlation with constant error, Test III, remains highly significant.

Table 2. *Pearson correlations between scores of four size estimation tests\**

( $N = 60.$ )

Mean score	Correlations between scores of tests					
	I:II	I:III	I:IV	II:III	II:IV	III:IV
Judgement time	0.78	0.69	—	0.77	—	—
Number of stops	0.63	0.67	—	0.66	—	—
Number of centrations on standard	0.71	0.75	—	0.80	—	—
Time on standard	0.61	0.66	—	0.77	—	—
Percentage of time on standard	0.54	0.46	—	0.76	—	—
Time per centration on standard	0.67	0.61	—	0.84	—	—
Redundant scanning time	0.73	0.75	—	0.70	—	—
Constant error	0.16	0.19	0.10	0.55	0.28	0.42
Constant error ( <i>r</i> 's with age partialled out)	—	—	0.06	—	0.25	0.42

\* 0.26 = 0.05 level; 0.33 = 0.01 level; 0.42 = 0.001 level.

Table 3. *Pearson correlations between pairs of logically independent scanning scores within size estimation Tests I, II and III\**

( $N = 60.$ )

Mean score	Test	Mean score					
		Number of centrations on standard	Time on standard	Time per centration on standard	Judgement time	Number of stops	Percentage of time on standard
Redundant scanning time	I	0.55	0.68	0.28	0.73	0.69	0.30
	II	0.62	0.61	0.24	0.58	0.61	0.28
	III	0.29	0.43	0.19	0.38	0.34	0.15
Number of stops	I	0.79	0.66	0.09	—	—	—
	II	0.72	0.70	0.27	—	—	—
	III	0.57	0.51	0.11	—	—	—
Percentage of time on standard	I	0.54	—	—	—	—	—
	II	0.44	—	—	—	—	—
	III	0.36	—	—	—	—	—

\* 0.26 = 0.05 level; 0.33 = 0.01 level; 0.42 = 0.001 level.

The results presented in Table 3 support the interpretation that these subjects manifested consistent scanning syndromes, rather than unrelated bits of scanning behaviour. Redundant scanning time, for example, is impressively related to judgement time. It is also apparent that subjects who looked back and forth many times before committing themselves to a size judgement spent greater percentages of time examining the standard stimuli (the number of centrations on the standard equalled those on the comparison stimulus).

Table 4 contains rank-order correlations of Test I scores with retest scores and Test V (40 mm. fixed black disk) scores for the twenty subjects seen again after a mean interval of 60 days. Although these correlations may be elevated somewhat by the fact that the twenty subjects retested consisted of extreme groups of extensive and limited scanners, it would appear that Test I scores are fairly reliable over these periods of time (Hypothesis 1). In the case of the constant error score, the retest correlation of 0.60 is similar to the  $r$  of 0.68 obtained in an independent study by Gardner & Long (1960c) after an interval of 3 years. The rho's between Test I scores and Test V scores generally support these indications of reliability and also demonstrate that the observed consistency of scanning pattern and constant error extends to judgements of a fixed black disk.

Table 4. *Rank-order correlations of Test I scores with retest and Test V scores\**

Mean score	Rho	
	Retest	Test V
Judgement time	0.71	0.77
Number of stops	0.78	0.91
Number of centrations on standard	0.67	0.68
Time on standard	0.60	0.75
Percentage of time on standard	0.66	0.65
Time per centration on standard	0.63	0.34
Redundant scanning time	0.81	0.83
Constant error	0.60	0.52

\* 0.44 = 0.05 level; 0.56 = 0.01 level; 0.68 = 0.001 level.

### (iii) *Duration of centration and apparent size*

Hypothesis 2, which stated (following Piaget) that the apparent size of a stimulus is a function of the duration of a single centration upon it, is confirmed by the analysis of variance (Linguist's Type I Mixed Design) summarized in Table 5. Mean constant errors for the four durations of centration are: 1 sec., -2.04 mm.; 2 sec., -1.86 mm.; 3 sec., -1.16 mm.; 4 sec., -1.11 mm. The effect was apparently approaching a maximum at 4 sec.

### (iv) *Individual consistencies in centration effects*

Hypothesis 3, which stated that individuals are consistently different in centration effects under identical viewing conditions, is apparently confirmed by the highly significant ( $P < 0.001$ ) Pearson  $r$ 's between the sixty subjects' constant errors for the four durations of centration in Test IV. These  $r$ 's range from 0.52 to 0.67, and

with age correlations partialled out, from 0.42 to 0.64. These  $r$ 's could be elevated, however, by possible individual consistencies in starting-position effects, since only ascending judgements were made.

Table 5. *Summary of analysis of variance of Test IV constant error scores*

(N = 60.)				
Source	D.F.	M.S.	F	P
Between subjects	59	15.504	—	—
Orders	3	24.847	1.656	N.S.
Error (b)	56	15.004	—	—
Within subjects	180	2.438	—	—
Duration of centration	3	13.687	6.138	< 0.001
Duration of centration $\times$ orders	9	2.573	1.154	N.S.
Error (w)	168	2.230	—	—
Total	239	—	—	—

(v) *Repeated judgements and apparent size*

Test III, in which subjects made ten ascending judgements of a 30 mm. light circle from the same 22 mm. starting-point, provided an opportunity to test the hypothesis that repeated viewing of a standard stimulus leads to an increase in its apparent magnitude (Hypothesis 4). No significant trend is present, however, in the mean constant errors for the ten trials of this test.

(vi) *Scanning and centration effect as determinants of apparent size*

Correlations between scanning and constant error scores are presented in Table 6. Hypothesis 5 appears to receive some limited confirmation in Test I, but not in Tests II and III. In Test I, extensive scanning, in terms of number of centrations on the standard, is associated with relatively little of the overestimation that characterized the subjects as a group (see Table 1) in this procedure. This finding suggests that extensive scanners are characterized by a consistent 'counteractive' approach (see Piaget *et al.* 1958) that tends to overcome centration effects and the other illusion-producing characteristics of the Test I disks.

Table 6. *Pearson correlations between scanning and constant error scores within three size estimation tests\**

Mean score	Correlations with constant error		
	Test I	Test II	Test III
Judgement time	-0.16	0.07	0.17
Number of stops	-0.29	0.04	0.00
Number of centrations on standard	-0.32	0.06	0.06
Time on standard	-0.19	0.14	0.17
Percentage of time on standard	-0.11	0.25	0.10
Time per centration on standard	0.10	0.21	0.21
Redundant scanning time	-0.19	-0.22	-0.08

\* 0.26 = 0.05 level.

(In the course of another recent study employing an independent group of thirty-three female subjects ranging in age from 21 to 25 years, it was possible to replicate Test I. The Pearson  $r$  between the mean number of centrations on the standard and



constant error was  $-0.32$ , identical to the correlation observed in the present study. These results seem to support the inference that extensive scanners are engaged in counteractive manoeuvres designed to overcome the illusion of enhanced magnitude produced by these hand-held disks. Further support for the relationship of number of centrations to constant error under the Test I judgement conditions derives from the fact that rho's between these scores were  $-0.32$  for the readministration of Test I and  $-0.35$  for Test V.)

Hypothesis 6 is also confirmed only partially. As noted earlier, centration effect independent of scanning (constant error, Test IV) is significantly correlated with constant error, Test III, and nearly significantly with constant error, Test II, but not with constant error, Test I. Since only ascending judgements were made in Tests II and III, however, the first two of these  $r$ 's could be elevated by possible individual consistencies in starting-position effects. None of the correlations between constant error, Test IV, and the scanning scores used in the present study is significant. Taken together, these results suggest that extensiveness of scanning is a significant determinant of apparent size only in the relatively complex and difficult Test I, whereas centration effect may be a significant determinant of apparent size only in the simpler and easier Test III. This interpretation is supported by the fact that multiple correlations (employing each scanning score with the centration effect score as predictors of apparent size in the free scanning tests) were in no case significantly higher than the correlations based on the individual predictors.

(vii) *Performances of extreme scanning groups in additional tests*

*Reversed-L Test.* As noted earlier, Hypothesis 7, which stated that extensiveness of scanning in size estimation tests is negatively related to error of the standard in judgements of inverted-T figures, could not be tested. The extreme scanning groups did differ significantly in the predicted direction, however, in errors of the standard with the unfamiliar reversed-L figures ( $t = 2.144$ ,  $P < 0.05$ ). The implication is that extensive scanners in size estimation tests are also extensive scanners in the reversed-L tests, with predictable consequences for error of the standard.

*Free Association Test.* The relationship between extensiveness of scanning and the percentage of responses to this test in the 'most distant' category, in terms of conceptual distance from the stimulus words, is significant ( $P < 0.05$ ), but in the opposite direction from that predicted (Hypothesis 8). (Because of the large proportion of extremely high or low percentage scores, White's non-parametric test was used.)

*Rorschach Test.* As noted earlier, the two judges agreed on a total of fourteen cases in each of three Rorschach ratings. Only these cases were used to test hypotheses concerning relations between extensiveness of scanning and isolation, projection, and delay. By Fisher's Exact Test, none of the relations between the three sets of ratings reaches significance, although isolation tends to be associated with projection ( $P = 0.12$ ) and delay ( $P = 0.17$ ).

As predicted (Hypotheses 9-11), significantly more Rorschach protocols of extensive scanners than limited scanners were judged high in indications of isolation, projection, and generalized delay ( $P = 0.01$  in each instance, by Fisher's Exact Test).

In keeping with Hypothesis 12, extensive scanners in size estimation tests spent significantly greater time than limited scanners surveying the inkblots before offering

first responses ( $t = 2.592$ ,  $P = 0.05$ ). Their slowness to respond to the inkblots thus seems analogous to their delay in committing themselves to size judgements, which took the form of extensive scanning of the standard and comparison stimuli. Also as predicted (Hypothesis 13), significantly more extensive scanners than limited scanners offered comments or responses in the Rorschach Test implying preoccupation with symmetry ( $P = 0.01$ , by Fisher's Exact Test). The two groups did not differ significantly in total number of responses or in any of the other major scoring categories employed by Rapaport *et al.* (1946). It should be noted here that the first of these specific Rorschach indicators is probably not independent of the ratings of generalized delay, and the second not independent of the ratings of isolation made by the judges.

### DISCUSSION

The present study demonstrates individual consistencies in attentional strategies that are apparent both in a variety of size estimation tests and in the Rorschach Test. The nature of the dimension of cognitive control called extensiveness of scanning, with its relationship to ratings of generalized delay, has implications for individual differences in judgement- or decision-time in a wider variety of situations. The essential difference between the performances of extensive and limited scanners in these procedures seems to lie in the amount of information sampled before commitment to a response. The extensive scanner seems preoccupied with the veridicality, exactness, and acceptability of his response. The limited scanner seems to accept his perceptual experience in a more relaxed and less critical way.

The confirmation of Piaget's hypothesis that the duration of a single centration is a determinant of apparent size has important implications for perception outside the laboratory. It is rare, for example, that one engages in extended comparisons of stimuli of the kind elicited by such laboratory procedures as the free scanning tests of the present study. Test IV, which involved single centrations upon a standard stimulus, may therefore approximate the normal conditions of perception more closely than the other size estimation procedures. The observed negative correlation of the centration effect with age was unanticipated and remains unexplained.

Our attempt to assess the relationship of scanning and centration effect to apparent size in free-scanning tests may have been affected by the judgement method employed. Each of the psychophysical methods introduces artifacts into judgements of apparent size. As employed in size estimation Tests I and II, for example, the method of adjustment rests upon the assumption that each subject's ascending and descending starting-position effects are equal. As employed in Tests III and IV (which contained only ascending judgements), this method presumably introduces unwanted individual differences in starting-position effects into apparent size scores. In future studies, Test IV should be expanded to include a wider range of centration times and to include both ascending and descending judgements, so that the apparent individual consistencies in centration effect can be assessed independent of starting-position effects.

The finding that extensive, rather than limited, scanners produced smaller percentages of 'most distant' responses to the free association test is at variance with anticipation and seems difficult to explain. Response to this test is a complex process involving a variety of variables, as indicated by its loading on several factors in the

study by Gardner *et al.* (1959). It could be that the relationship observed in the present study is determined by unassessed variables related to scanning, rather than to scanning itself. Further studies are needed to clarify this unanticipated result.

It is theoretically important, but not surprising, that extensive scanners were rated higher than limited scanners in employment of the defence mechanisms of isolation and projection. Both defences imply concern with exactness. In the case of the extensive scanner who employs isolation to a great degree, careful examination of stimuli before making decisions is apparently associated with doubt and equivocation. The extensive scanner in whom projection is relatively strong may show concern with exactness that is an accompaniment of suspicion and distrust. Some extensive scanners, however, were rated low in both isolation and projection. Further and more extensive research is required to determine the full pattern of relationship between this aspect of attention deployment and defence organization.

### REFERENCES

- FORD, A. & LEONARD, J. L. (1958). Techniques for recording surface bioelectric direct currents. *Research Report, U.S. Navy Electronic Laboratory, San Diego, California*, no. 839.
- FORD, A., WHITE, C. T. & LICHTENSTEIN, M. (1959). Analysis of eye movements during free search. *J. Opt. Soc. Amer.* **49**, 287-92.
- FREUD, S. (1922). Certain neurotic mechanisms in jealousy, paranoia and homosexuality. In *Collected Papers*, vol. II. London: Hogarth Press, 1948.
- FREUD, S. (1926). *The Problem of Anxiety*. New York: Norton, 1936.
- GARDNER, R. W. (1961). Cognitive controls of attention deployment as determinants of visual illusions. *J. Abnorm. Soc. Psychol.* **62**, 120-7.
- GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, H. B. & SPENCE, D. P. (1959). Cognitive control: a study of individual consistencies in cognitive behavior. *Psychol. Issues*, **1**, no. 4.
- GARDNER, R. W. & LONG, R. I. (1960a). Errors of the standard and illusion effects with the inverted-T. *Percept. Mot. Skills*, **10**, 47-54.
- GARDNER, R. W. & LONG, R. I. (1960b). Errors of the standard and illusion effects with L-shaped figures. *Percept. Mot. Skills*, **10**, 107-9.
- GARDNER, R. W. & LONG, R. I. (1960c). The stability of cognitive controls. *J. Abnorm. Soc. Psychol.* **61**, 485-7.
- KRIS, C. (1960). Vision: Electro-oculography. In *Medical Physics*, vol. III. Ed. Glasser, O. Chicago: Yearbook.
- PIAGET, J. (1947). *The Psychology of Intelligence*. London: Routledge and Kegan Paul, 1950.
- PIAGET, J. (1961). *Les Mécanismes Perceptifs*. Paris: Presses Universitaires de France.
- PIAGET, J., VINH-BANG, & MATALON, B. (1958). Note on the law of the temporal maximum of some optico-geometric illusions. *Amer. J. Psychol.* **71**, 277-82.
- PIAGET, J. *et al.* (1942-56, 1959). Recherches sur le Développement des Perceptions. Recherches I-XXXII, XXXVII, XXXVIII. *Arch. Psychol., Geneve*, **29-35**, **37**.
- RAPAPORT, D., GILL, M. & SCHAFER, R. (1946). *Diagnostic Psychological Testing*, vol. II. Chicago: Yearbook.
- SCHAFER, R. (1954). *Psychoanalytic Interpretation in Rorschach Testing*. New York: Grune & Stratton.
- SHACKEL, B. (1960). Pilot study in electro-oculography. *Brit. J. Ophthalm.* **44**, 89-113.
- SHACKEL, B. & DAVIS, J. R. (1960). A second survey with electro-oculography. *Brit. J. Ophthalm.* **44**, 337-46.

(Manuscript received 28 March 1961)



## STIMULUS WAVELENGTH VARIATION AND SIZE AND DISTANCE JUDGEMENTS

By RAY OVER

*Psychology Department, Queen's University, Belfast*

Previous experiments which have studied the relationship between stimulus wavelength variation and size and distance judgements have confounded wavelength with luminance, which is known to be an effective variable. Using a rationale which combines the fact that light from stimuli of different wavelengths stimulates different areas of the retina (chromatic aberration) with the finding that under reduced viewing conditions two stimuli of the same wavelength are judged equal in both size and distance when they stimulate equal areas of the retina, it was predicted that when the subject was presented with two stimuli subtending the same visual angle and of the same luminance but of different wavelengths the stimulus of longer wavelength would be judged to be both larger and closer than the stimulus of shorter wavelength. This prediction was confirmed in the experiment.

### I. INTRODUCTION

The present aim is to determine whether stimulus wavelength is a variable which affects size and distance judgements. If two stimuli are judged equal in both size and distance when they are of the same wavelength, it will be asked whether changes in the wavelength of one of the stimuli will be followed by judgements that the stimuli now are not equal in size and distance.

The existing experimental evidence is conflicting. Three experiments have investigated the relationship between stimulus 'colour' and size judgements given under unrestricted viewing conditions. Warden & Flynn (1926) instructed subjects to rank in terms of apparent size eight objects of the same physical size but differing in hue. They found no relationship between size rankings and spectral values. Gunlach & Macoubray (1931), repeating this experiment, found that objects were consistently ranked from largest to smallest in the order yellow, white, green, blue, red, grey, and black. With a paired comparison technique they found the order of ranking to be green, yellow, white, blue, red, grey, and black. Wallis (1935), also using a paired comparison method, found that yellow was judged largest, followed by white, red, green, blue, and black. In the latter two experiments a high positive correlation was reported between the luminous intensities of the stimulus and size rankings. It is known that brighter objects are judged to be larger than less bright objects (Robinson, 1954). Accordingly it is possible that the above results, and the differences between them, are artifacts of different experimental confoundings of stimulus wavelength and luminance. It has not yet been demonstrated that variation of stimulus wavelength, independent of variation in the luminance of the stimulus, influences size judgements.

Similar differences in results have been seen in those experiments which have studied the relationship between stimulus hue and distance judgements. Using unrestricted viewing conditions Taylor & Sumner (1945) and Johns & Sumner (1948) measured the distances at which a grey comparison stimulus had to be placed in order that it should be judged equal in distance to each of seven coloured stimuli set at a

constant distance from the subject. Although no relationship was found between comparison distances and spectral values, a high positive correlation was found between comparison distances and luminous intensities. It is known that variation in the luminous intensity of a stimulus affects distance judgements (Coules, 1955). Under reduced viewing conditions with the two stimuli at the same physical distance from the subject, Luckiesh (1918) noted that subjects judged the red stimulus to be closer than the blue, while Pillsbury & Schaeffer (1937) found that the blue stimulus was judged closer than the red. The luminous intensities of the stimuli are not reported in these last two experiments; it is possible that the differences in the results are a function of different confoundings of stimulus wavelength and luminance. Although Karwoski & Lloyd (1951) have shown that judgements of the distance of blue objects are less accurate than judgements of the distance of red objects, no experiment appears to have been reported showing that a stimulus of one wavelength is judged to be closer than a stimulus of another wavelength when the two stimuli are of the same luminous intensity.

A rationale can be developed to relate stimulus wavelength and size and distance judgements under limiting experimental conditions. The area of the retina stimulated is a function of the wavelength of the stimulus. Duke-Elder (1942, p. 1079) states that, '...since blue rays are refracted more than red rays by the ocular media, their foci not only lie at different levels (chromatic aberration) but make different angles with the optic axis and will thus stimulate disparate points'. If two stimuli which differ in wavelength but subtend the same visual angle at the nodal point of the eye are viewed along the line of normal fixation, the stimulus of longer wavelength will stimulate a larger area of the retina than the stimulus of shorter wavelength (Duke-Elder, 1954, p. 52). Under reduced viewing conditions two stimuli are judged equal in both size and distance only when they stimulate equal areas of the retina (Over, 1960). It would thus be predicted under these viewing conditions that, when two stimuli differ only in wavelength, the stimulus of longer wavelength would be judged to be both larger and closer to the subject than the stimulus of shorter wavelength, and that it would be necessary to change the visual angle subtended by the stimuli until both stimuli stimulate equal areas of the retina before judgements of size and distance equality will be given. It is difficult, on the other hand, to determine the relationship between stimulus wavelength and size and distance judgements under unrestricted viewing conditions, for under these viewing conditions both size equality judgements (Gilinsky, 1955) and distance equality judgements (Over, 1961) are independent of the relative areas of the retina stimulated by the two objects.

In the experiment to be reported the subjects, tested under reduced viewing conditions, were instructed to judge the relative sizes and distances of two stimuli of the same luminous intensity but differing in wavelength. It was predicted that when the two stimuli subtended the same visual angle, the stimulus of longer wavelength would be judged to be both larger and closer than the stimulus of shorter wavelength.

## II. EXPERIMENTAL

### (i) *Apparatus*

The experimental situation was similar to the reduced viewing conditions under which two objects are judged equal in both size and distance when they subtend the same visual angle (Over, 1960). Testing was carried out in an almost totally dark room 20 ft. long. The subject was seated at a desk inside a booth with a headrest 3 ft. from the floor. A black curtain at the front of the booth could be raised or lowered so that when it was lowered the subject could see no light and when it was raised he could see only the two stimulus objects. Stray light from the walls and floor was trapped by screens 18 in. in front of the headrest.

The standard and comparison stimuli were produced by the illumination of filters, mounted on cardboard covers, which were placed on a milk-glass screen at the front of two identical boxes. Each of the 3 in. square filters was illuminated by a 15 W. 230 V. lamp from within the box and an iris between the lamp and filter could be adjusted to allow change in the luminous intensity of the filter without change in the temperature of the lamp filament. Three pairs of filters were used. These were gelatin Wratten filters Nos. 29, 47 B, and 61, which transmit radiation between 600 and 700, 400 and 500, and 500 and 600 m $\mu$ , respectively, and which have the phenomenal hues red, blue, and green respectively. The luminous intensities of all the filters were kept constant at 0.0447 ft. L., as measured by an S.E.I. photometer. When the curtain of the booth was lifted the subject had the impression that there were two objects at eye-level out in space in front of him and that these were equal in brightness and either similar or different in hue (depending on the combination of filters used).

### (ii) *Procedure*

Two groups, each of 12 subjects, were used. One group made size judgements, the other distance judgements. The majority of the subjects were undergraduate students and were not aware of the aim of the experiment. All reported having normal vision.

Subjects who made size judgements were instructed:

When the curtain is lifted you are to judge whether the object on your right is larger, smaller, or equal in size to the object on your left, when by equal in size is meant that if you were to walk out to the objects in turn and place a tapemeasure across their horizontal diagonals, you would obtain exactly the same reading.

Subjects who made distance judgements were told:

When the curtain is lifted you are to judge whether the object on your right is farther, closer, or the same distance from you as the object on your left when by the same distance is meant that if you were to place a tapemeasure between yourself and the objects in turn you would obtain the same reading.

Subjects were tested individually and a dark adaptation period of 4 min. preceded testing. Viewing was monocular with the right eye. The three pairs of filters allowed nine combinations of standard stimulus wavelength and comparison stimulus wavelength; for three arrangements the pair of stimuli were the same hue, and for the other six they were of different hues. Each subject made sets of judgements for all nine combinations, the order of presentation differing between subjects. For half of the subjects of each group the standard stimulus was on the subject's left and the comparison on the right, and for the other half the standard was on the right and the comparison on the left.

The two stimuli were initially 10 ft. from the subject and thus subtended the same visual angle. The standard always remained 10 ft. from the subject but the comparison could be moved by 15 in. steps along a track towards or away from the subject so that the centres of the two stimuli were always separated by 9° with respect



to the subject. Each subject made judgements for a particular combination of stimulus wavelengths when the comparison stimulus was 10 ft., 11 ft. 3 in., and 8 ft. 9 in. away, the order of presentation of the two latter positions being determined at random. If judgements of 'smaller' or 'farther' were given with the second of the above positions, of 'larger' or 'closer' with the third, the comparison stimulus was moved by further 15 in. steps until the judgement was reversed. With the stimulus constant in size and luminous intensity, changes in distance change the total amount of light subtended at the eye by the stimulus but not the amount of light subtended per unit solid angle. Even if the latter, rather than the former, is an effective variable in size and distance judgements, it could operate only to restrict the expected trends and cannot be considered to produce them.

It can be seen that as the stimuli were of constant physical size, changes in the visual angle subtended by the comparison stimulus were brought about by changes in the distance of the comparison stimulus. Both size and distance judgements can justifiably be related to such variation because under reduced viewing conditions size judgements are a function of variation as much in the distance as in the size of the stimulus; distance judgements as much a function of variation in the size as in the distance of the stimulus (Over, 1960). When questioned after the experiment all subjects who made size judgements considered that inter-trial differences between the visual angles subtended by the comparison stimulus reflected differences in physical size, while all subjects who made distance judgements considered that the differences reflected changes in physical distance.

### III. RESULTS AND DISCUSSION

The mean visual angles subtended by the comparison stimulus when it was judged equal in size and in distance to the standard stimulus are set out in Table 1 for the various combinations of standard and comparison wavelengths. In compiling these results, judgements that the comparison stimulus was larger (closer) than the standard when the comparison stimulus subtended a certain visual angle, but that the comparison stimulus was smaller (farther away) when it subtended a smaller visual angle were considered equality judgements at the intermediate visual angle. The results show that in general two stimuli of the same wavelength are judged equal in both size and distance when they subtend the same visual angle, while stimuli of longer wavelengths are judged equal to stimuli of shorter wavelengths only when the former subtend smaller visual angles than the latter. From the analysis of variance summarized in Table 2 it can be seen that the differences in the response measures brought about by variation in stimulus wavelength are significant. The fact that differences between the size judgement visual angle and the distance judgement visual angle values are insignificant, together with the insignificant interactions, indicates that stimulus wavelength variation affects size and distance judgements in a similar way.

The dispersion of values around the means is reported in Table 3. The standard deviations are approximately the same as those found in an earlier experiment in which stimulus wavelength was not an experimental variable (Over, 1960). Although this has not been tested, it is thought that variation between subjects would be

equal to variation between the various judgements made by a particular subject on repeated presentation of a stimulus arrangement.

It may be possible to account for the results in terms of an associative, rather than a retinal, mechanism. Johns & Sumner (1948, p. 28) suggest, for example, that blue objects are judged to be farther away than red objects because of 'past experience

Table 1. *Mean visual angles subtended by comparison stimuli*

Comparison hue	Standard hue		
	Red	Green	Blue
(a) For size equality judgements			
Red	1° 56'	1° 42'	1° 41'
Green	2° 13'	1° 59'	1° 47'
Blue	2° 24'	2° 10'	1° 58'
(b) For distance equality judgements			
Red	1° 56'	1° 45'	1° 29'
Green	2° 29'	2° 04'	1° 43'
Blue	2° 39'	2° 24'	2° 03'

Visual angle of standard = 2° 02'.

Table 2. *Analysis of variance summary*

Source	S.S.	D.F.	V.E.	F	P
1. <i>A</i>	138.12	2	64.06	31.56	0.001
2. <i>B</i>	164.51	2	82.26	40.52	0.001
3. <i>C</i>	1.34	1	1.34	0.66	—
4. <i>AB</i>	3.63	4	0.91	0.44	—
5. <i>AC</i>	9.01	2	4.51	2.22	—
6. <i>BC</i>	7.95	2	3.98	1.96	—
7. <i>ABC</i>	2.74	4	0.69	0.34	—
8. Within	401.93	198	2.03	—	—
Total	729.22	215			

*A* is the hue of the standard (red, green, blue); *B* the hue of the comparison (red, green, blue); and *C* the response measure (size judgements, distance judgements).

Table 3. *Dispersion of values around the mean comparison visual angles reported in Table 1*

Comparison hue	Standard hue		
	Red	Green	Blue
(a) For size equality judgements			
Red	11' 02"	10' 24"	16' 24"
Green	12' 24"	6' 43"	11' 36"
Blue	11' 02"	6' 43"	10' 43"
(b) For distance equality judgements			
Red	6' 43"	11' 36"	17' 31"
Green	8' 00"	7' 07"	10' 24"
Blue	10' 24"	8' 24"	9' 46"

Standard deviations in minutes and seconds of visual angle.

with aerial perspective'. There are, too, various unformalized notions that red is a more 'warm', 'friendly', etc., colour than blue and that 'warm' objects are judged to be closer than 'cold'. In the present experiment the explicit judgemental criteria, which expressed size and distance equality in terms independent of wavelength, would have acted against the operation of any associative mechanism. It was stated earlier that if retinal relationships alone determine the effect of stimulus wavelength variation on size and distance judgements, both size and distance judgements given under unrestricted viewing conditions should be independent of stimulus wavelength variation (size- and distance-constancy). If an associative mechanism is responsible for the results reported in the above experiment, it would be expected that results following the same relationship would be found under unrestricted viewing conditions. No adequate data, obtained with stimuli of the same luminous intensity, are available to determine this issue.

The results also bear on that form of the size-distance invariance hypothesis which states that there are experimental features other than the physical sizes and distances of the stimuli, the variation of which will lead to changes in both size and distance judgements such that judgements of 'larger' will be accompanied by judgements of 'closer' and not by judgements of 'farther'. Experiments by Ittelson & Kilpatrick (1952), Gruber (1954), and Rump (1961) suggest that this relationship does not hold, but there is some doubt whether in these experiments both the size and distance judgements were obtained under equivalent experimental conditions (e.g. with size and distance instructions which emphasized the same judgemental criteria). The present experiment, together with the experiments by Coules (1955) and Robinson (1954) which found that variation in the luminous intensities of the stimuli produced similar size and distance judgement functions, satisfies this criterion. Further studies of the effect of stimulus variation on *both* size and distance judgements are required.

## REFERENCES

- COULES, J. (1955). Effect of photometric brightness on judgments of distance. *J. Exp. Psychol.* **50**, 19-25.
- DUKE-ELDER, W. S. (1942). *Text-book of Ophthalmology*, vol. 1. London: Henry Kimpton.
- DUKE-ELDER, W. S. (1954). *The Practice of Refraction*. London: Churchill.
- GILINSKY, A. S. (1955). The effect of attitude upon the perception of size. *Amer. J. Psychol.* **68**, 173-92.
- GRUBER, H. E. (1954). The relation of perceived size to perceived distance. *Amer. J. Psychol.* **67**, 411-26.
- GUNLACH, C. & MACOUBRAY, C. (1931). The effect of color on apparent size. *Amer. J. Psychol.* **43**, 109-11.
- ITTELSON, W. H. & KILPATRICK, F. P. (1952). The size-distance invariance hypothesis. In Kilpatrick, F. P. (ed.), *Human Behavior from the Transactional Point of View*. New Hampshire: Institute for Associated Research.
- JOHNS, E. H. & SUMNER, F. C. (1948). Relation of the brightness differences of colors to their apparent distances. *J. Psychol.* **26**, 25-9.
- KARWOSKI, T. F. & LLOYD, V. V. (1951). Studies in vision: V. The role of chromatic aberration in depth perception. *J. Gen. Psychol.* **44**, 159-73.
- LUCKIESH, M. (1918). On 'retiring' and 'advancing' colors. *Amer. J. Psychol.* **29**, 182-6.
- OVER, R. (1960). Size and distance judgements under reduction conditions. *Aust. J. Psychol.* **12**, 162-8.
- OVER, R. (1961). Distance constancy. *Amer. J. Psychol.* **74**, 308-10.



- PILLSBURY, W. B. & SCHAEFFER, B. R. (1937). A note on 'advancing' and 'retreating' colors. *Amer. J. Psychol.* **49**, 126-30.
- ROBINSON, E. J. (1954). The influence of photometric brightness on judgments of size. *Amer. J. Psychol.* **67**, 464-74.
- RUMP, E. E. (1961). The relationship between perceived size and perceived distance. *Brit. J. Psychol.* **52**, 111-24.
- TAYLOR, I. & SUMNER, F. C. (1945). Actual brightness and distance of individual colors when their apparent distance is held constant. *J. Psychol.* **19**, 79-85.
- WALLIS, W. A. (1935). The influence of color on apparent size. *J. Gen. Psychol.* **13**, 193-9.
- WARDEN, C. J. & FLYNN, E. L. (1926). The effect of color on apparent size. *Amer. J. Psychol.* **37**, 398-401.

(*Manuscript received 30 May 1961*)



## DISINHIBITION AND THE REMINISCENCE EFFECT IN A MOTOR LEARNING TASK

By S. RACHMAN

*Institute of Psychiatry (Maudsley Hospital), University of London*

This study was designed to investigate the operation of disinhibition in humans. Two groups of subjects were made to perform a motor learning task (pursuit rotor). The subjects in the experimental group were presented with a brief disinhibiting, alien stimulus during their 5 min. practice period. It was predicted that the introduction of this alien stimulus would produce an augmentation of performance level and a reduction in the reminiscence effect. Both predictions received partial confirmation and the results are interpreted in terms of Pavlov's description of inhibition and Eysenck's account of the development and dissipation of reactive inhibition. In an attempt to clarify some aspects of the results, a second experiment was carried out. Four groups of ten subjects each were given the same pursuit rotor task to complete. Group A acted as a control, group B was presented with an alien stimulus early in the 5 min. practice period, group C late in the practice period and group D very late in the practice period. In addition to clarifying some of the earlier results this experiment showed that the effect of an alien stimulus on performance is most marked when it is introduced late in the practice period.

### INTRODUCTION

In a series of unpublished experiments on psychomotor performance a negative reminiscence effect was unexpectedly obtained on two separate occasions. In both cases, this effect appeared after the introduction of conflicting stimuli. It was suggested that the probable reason for the non-appearance of reminiscence was that the conflict had acted as a disinhibitor. It was felt, however, that this explanation required some form of external confirmation. The present experiment then was designed in an attempt to demonstrate that the deliberate introduction of an alien stimulus during practice can interfere with the normal course of the inhibitory process and produce a negative reminiscence effect.

The concept of disinhibition was introduced by Pavlov (1927) who stated that '...the temporary restoration of the reflex which is in the process of extinction, or which is already extinguished, (is) based upon the removal of an inhibitory process. We can describe this phenomenon as a dis-inhibition, a term which we shall always use in the future when we wish to denote a temporary removal of inhibition.' Pavlov emphasized the temporary nature of disinhibition and claimed that it was produced by the introduction of novel stimulation. 'The removal of inhibition...which is affected under the influence of any alien stimulus...is only temporary' (Pavlov, 1927). Although experimental demonstrations of disinhibition in humans are comparatively rare, the available evidence (Hovland, 1936; Razran, 1939) supports Pavlov's account. Disinhibition can be produced by introducing alien stimuli and it is temporary in nature.

As there is a considerable amount of information available concerning reminiscence effects on the pursuit rotor (Ammons, 1955; Eysenck, 1956; Jones, 1960; Treadwell, 1960) this motor task was chosen for the present experiment. A practice-rest pattern which has regularly produced positive reminiscence is that of 5 min.



massed practice followed by 10 min. of rest and then a further 1 min. of practice (Jones, 1960). Consequently, this pattern of work and rest was used here.

The prediction used in this experiment was that *the introduction of an alien or extraneous stimulus late in the practice period will produce disinhibition and a decreased reminiscence effect.*

## EXPERIMENT I

### Method

A pursuit rotor with a 10 in. diameter disk rotating in a clockwise direction at 60 r.p.m. and with a  $\frac{3}{4}$  in. diameter contact point was used. The contact point or 'target' was set with its centre  $3\frac{1}{4}$  in. from the centre of the turntable and flush with the surface of the turntable. The stylus (total weight 2 oz.) consisted of a circular plastic handle  $4\frac{1}{2}$  in. long with a guard set 1 in. from the end of the handle. The extension rod (6 in. long,  $\frac{1}{16}$  in. diameter) was hinged so that only its weight rested on the turntable and it had an  $85^\circ$  bend, 1 in. from its end. When steady contact was made between the stylus and the target, the electric circuit was completed.

The two electric timers connected to the rotor were calibrated to read the time on target in tenths of a second. They were wired in such a way that they alternated in recording the time on target every ten seconds (the length of each trial). Calibration and zero adjustments were made prior to the testing of each subject and the entire apparatus switched on  $\frac{1}{2}$  hr. before testing commenced and kept running during all rest periods. The stylus point and target were cleaned with smooth emery-cloth after every 15 min. of operation.

The rotor was set on a table and the target was 31 in. off the floor. The recording clocks and scoring sheets were positioned so that the subject was unable to see them.

Two groups of subjects were used. Both groups were given 5 min. of continuous practice (thirty trials) followed by a 10 min. rest period. They were then re-tested for a further 1 min. In the case of the experimental group, however, a buzzer was sounded for 2 sec. during the continuous practice period. The buzzer was introduced towards the end of the practice session—after 4 min. 35 sec.

The general procedure was as follows:

The subject was shown the apparatus and instructed to attempt to keep the stylus on the target continuously, and to chase it when he had lost contact. After a demonstration by the experimenter, the subject was then given a brief trial on the apparatus to ensure that he comprehended the nature of the task. The subject was then instructed to proceed with the task until told to stop.

The subjects used consisted of ten males and five females in each group. The mean age of each group was: controls 26.4 (s.d. 7.6), experimental group 24.9 (s.d. 8.2).

The subjects were randomly assigned to the control or experimental groups. After completing the motor task, they were requested to complete the Maudsley Personality Inventory (Eysenck, 1959). The purpose of this procedure was to ascertain that the groups were not weighted in either the introvert or extravert direction, Eysenck (1956) having pointed out the correlation between reminiscence and extraversion. The mean extraversion scores for each group were: control 18.9 (s.d. 6.1); experimental 19.8 (s.d. 5.4). These means do not differ significantly ( $t = 1.3845$ ).

## RESULTS

The effect of the alien stimulus was measured by comparing pre- and post-buzzer performance levels for each group. As the buzzer was presented during trial 28, the difference in performance between trials 29 and 27 was calculated for each group. The control group showed a marked decline in performance at this stage as a result of accumulating inhibition, whereas the experimental group maintained its level of performance (see Fig. 1). The control group showed a mean decline of  $-0.46$  and the experimental group showed virtually no change with a mean of  $-0.02$ . This mean difference is significant at the 1% level ( $t = 3.3846$ ).

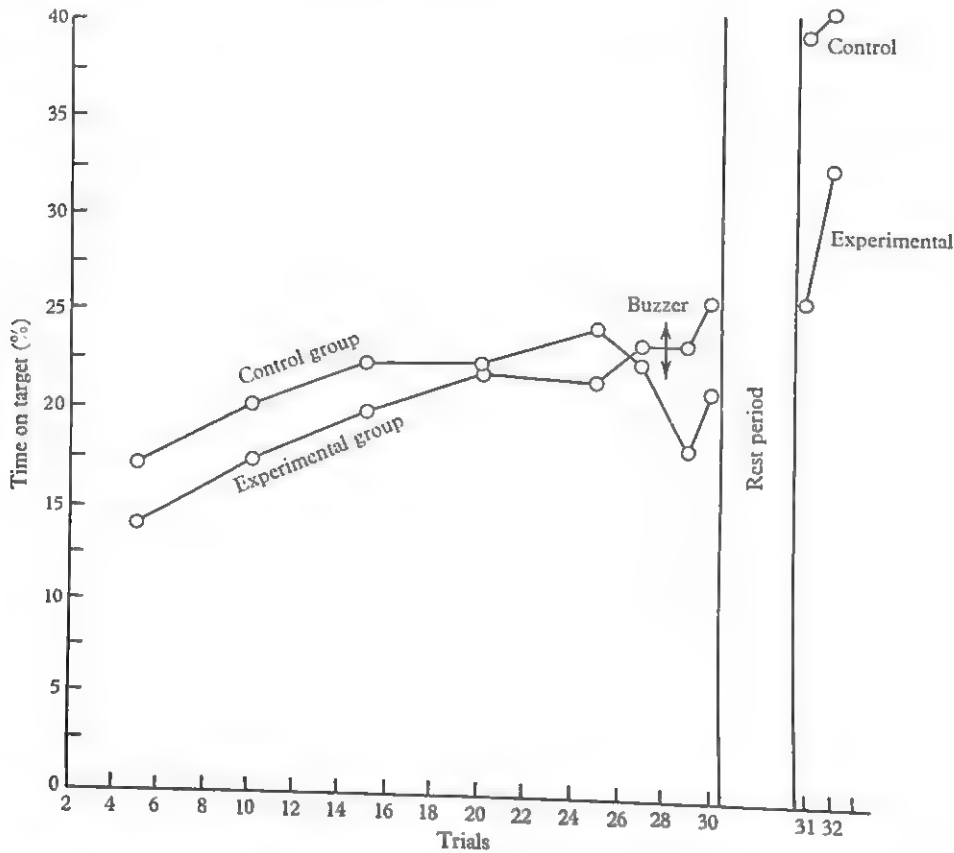


Fig. 1. Performance on a motor learning task measured in seconds on target. The rest period is 5 min. in duration.

The reminiscence scores for each subject were obtained by subtracting his score (i.e. time on target) on the last trial of the 5 min. practice period from his first post-trial. The mean reminiscence score for the control group was 1.42 (s.d. 1.11) and 0.29 (s.d. 0.86) for the experimental group (see Fig. 1). These means also differ at the 1% level ( $t = 2.866$ ).

## DISCUSSION

The introduction of an alien stimulus towards the end of the practice period was expected to produce a temporary augmentation of the response level. This augmentation should appear as a result of the disinhibiting effect of the buzzer. Because of this temporary increment in the performance which occurs just before the rest period, the experimental group should then produce a smaller reminiscence score than the control group.

Both of these predictions were borne out in the present experiment although the expected 'increment' in performance of the experimental group is only deduced by comparison with the control group's decline in performance. The control group showed a decline in performance towards the end of the practice period whereas the experimental group maintained its comparatively high level of performance. Secondly, the control group produced significantly greater reminiscence scores than the experimental group.

The disinhibitory mechanism postulated here can be accounted for in terms of Eysenck's (Eysenck, 1956, 1957; Eysenck & Maxwell, 1961) description of the development and dissipation of reactive inhibition in this task. Reactive inhibition accumulates during the performance of the pursuit task and is present in quantity towards the end of the 5 min. massed practice period. For purposes of interpretation, let us assume that the effect of the buzzer noise presented after 4 min. and 35 sec. of practice is to produce an involuntary rest pause. During this rest pause some of the  $I_R$  will dissipate. The noise produces a rest pause which lasts longer than the normally occurring involuntary rest pauses and in this way the experimental group's reactive inhibition undergoes more dissipation than that of the control group. The inhibitory potential of the experimental group is temporarily reduced or, to use Pavlov's term, disinhibited. Pavlov of course investigated the disinhibition of responses while the present study is concerned with the disinhibition of reactive inhibition. These two approaches are however compatible and it may even be possible to incorporate Pavlov's work into the broader framework of the modern two-factor theory of inhibition followed by the present writer.

This analysis of the effects of a disinhibiting stimulus on performance can probably be applied with profit to tasks other than the pursuit rotor. In discussing performance on vigilance tests, for example, Broadbent (1953) says that in continuous, monotonous tasks 'the occurrence of an unusual stimulus, the signal, should then improve performance and it will be recalled that the occurrence of a signal on the Dials partly reinstated the type of performance appearing early in the watch'. He points out further that by using auditory stimulation it is possible to reduce decrements in performance of vigilance-type tasks. Similar findings are reported by Fraser (1953).

The interpretation of the present experimental findings is however complicated by one unexpected aspect of the results. After both groups have dissipated the reactive inhibition during the 10 min. rest period they should again produce similar performances. In fact, the experimental group showed no improvement in performance after the rest period in contrast to the considerable gains made by the control group. The reason for this failure to improve is not easily explicable. As the reminiscence effects observed in pursuit rotor experiments are generally agreed to be a reflexion



of the dissipation of reactive inhibition during the rest period, a complete failure to recover after rest presumably indicates some disruption of the normal process of  $I_R$  dissipation.

#### SUMMARY

Two groups of fifteen subjects each were given 5 min. of massed practice on the pursuit rotor. After a 10 min. rest period they were re-tested and reminiscence scores derived for each subject. Towards the end of the practice period the experimental subjects were presented with a buzzer sound which acted as a disinhibitor. The effect of the buzzer was to maintain the performance of the experimental group whereas that of the control group deteriorated. It was also found that the disinhibited group produced significantly smaller reminiscence scores than the control group.

#### EXPERIMENT II

This experiment was designed with two purposes in mind. First, it was hoped to clarify the results obtained in the previous experiment on this topic. The second aim was to investigate the effect of introducing the 'disinhibitor' at different stages of the practice period.

More reactive inhibition is present at the latter part of the practice period than in the early stages. Hence, one would expect a disinhibitor to produce a greater effect if introduced towards the end of the task. A disinhibitor introduced at an early stage should produce a slight effect or no effect at all. These predictions may be stated in a formal manner:

- (1) An alien stimulus will produce disinhibition if it is introduced late in the practice period.
- (2) An alien stimulus introduced early in the practice period will produce little or no disinhibition.

The introduction of an alien stimulus produces an involuntary rest pause. During this pause some reactive inhibition dissipates. Hence, if the rest pause occurs soon before the formal rest period is taken, there will be less  $I_R$  to dissipate during this rest. This effect should then result in a reduced reminiscence score.

If the alien stimulus is introduced early in the practice period, however, the reminiscence score should not be noticeably effected. There are two reasons for this prediction. First, an alien stimulus presented early should produce little disinhibition. Secondly, even if some disinhibition were produced at an early stage, the subject should still accumulate a further substantial amount of  $I_R$  in the remaining period of practice. After a formal rest period is taken, the reminiscence effect should resemble that obtained under control conditions. The early introduction of an alien stimulus should at most result in a reminiscence score slightly below that of a control group. These predictions may be stated as follows:

- (3) The introduction of an alien stimulus late in the practice period will produce a decreased reminiscence effect; the introduction of an alien stimulus early in the practice period will produce little or no depression of the reminiscence score.

*Procedure*

As in the previous experiment, the pursuit rotor task was employed. The practice period consisted of thirty 10 sec. trials of massed practice. The rest period lasted, as before, for 10 min. After resting, each subject was then given three further 10 sec. trials.

Four experimental conditions were used. The control group A practised for 5 min. without interruption; group B subjects were presented with a 2-sec. buzzer sound after 1 min. of practice (early condition); group C (late condition) buzzed after 4 min. 35 sec. practice; group D subjects were buzzed after 4 min. 45 sec. practice (very late condition).

The *sample* consisted of forty normal subjects divided into four groups of ten subjects each. The groups were constituted as shown in Table 1. There were four females in each group.

Table 1

Group	<i>E</i>	S.D.	<i>N</i>	S.D.	Age	S.D.
A. Control	28.6	9.1	24.2	9.4	27.4	7.4
B. Early buzz	29.2	8.1	22.8	7.8	24.7	5.4
C. Late buzz	26.8	7.6	27.6	10.1	26.3	6.2
D. Very late buzz	26.8	10.4	25.0	8.6	26.4	5.7

*Results*

The effect of the alien (buzzer) stimulus was calculated by comparing the percentage time on target on the pre- and post-buzzer trials (e.g. score on trial 30-trial 28 in group D). The scores obtained from each of the three experimental groups in this manner were then compared with the relevant performance changes recorded by the control group on corresponding trials.

The results shown in Table 2 were obtained.

Table 2. *Mean post-buzzer increments*

Group	Trials		
	(7-6)	(29-27)	(30-28)
A (control)	-0.2	-0.7	+0.6
B	1.8	—	—
C	—	5.2	—
D	—	—	3.4

Group B, which received the buzzer after 1 min. of practice, showed a slight increment in performance (see Fig. 2) but this change was not significantly different from that recorded by the control group at the same point ( $t = 1.1462$ ).

Group C, which received the buzzer after 4 min. 35 sec. of practice, showed an increase in performance which was significantly greater than that of the control group at the same point ( $t = 2.2614$ ,  $P = 0.05$ ). Group D, which received the buzzer after 4 min. 45 sec. of practice, showed a greater increase in performance than the control group but this difference failed to reach the 5% level of significance ( $t = 1.9627$ ).

Reminiscence scores were calculated as before. The mean reminiscence scores for the four groups were: A (control) 18.6, B 11.1, C 4.7, D 5.8.

An analysis of variance was carried out on these results (Table 3).

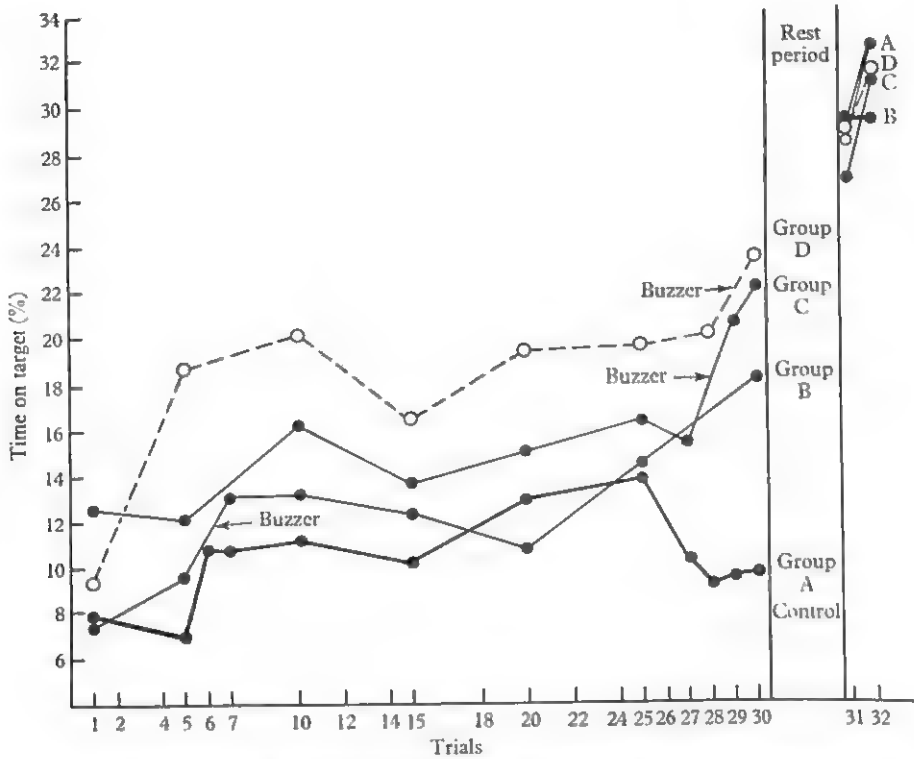


Fig. 2. Disinhibition and reminiscence in a motor learning task: temporal factors.

Table 3

Source	S.S.	D.F.	MSI	F	P
Between groups	2208.4	3	736.13	12.94	0.01
Within groups	2047.0	36	56.86	—	—
Total	4255.4	39	—	—	—

The control group's reminiscence scores were significantly greater than those of group B ( $P = 0.05$ ) and groups C and D ( $P = 0.01$ ). The reminiscence scores of groups B, C and D did not differ significantly.

### DISCUSSION

The results of the present experiment (see Fig. 2) sustain, with slight exceptions, the three hypotheses.

Hypothesis I, which states that the introduction of an alien stimulus late in the practice period produces disinhibition, is supported by the results of group C and the trend of group D's results is also in the predicted direction. Hypothesis II, which states that an alien stimulus presented early in the practice period will produce little or no disinhibition, is sustained by the results obtained with group B. The



group evidenced a slight increment in performance after the introduction of the buzzer but this change did not differ significantly from the control group's performance. Hypothesis III is borne out by the experimental results. The late-buzzer groups (C and D) showed significantly reduced reminiscence scores; the early-buzzer group (B) also produced a depressed reminiscence score but this effect was not as marked as was the case with groups C and D.

An aspect of the experimental results which warrants discussion is the performance of the control group. As can be seen from Fig. 2, the actual performance of the control group was poorer than that of the other three groups almost throughout. This fact should not influence the interpretation of the experiment unduly however, as we are concerned here with changes in performance rather than pursuit-rotor performance as such. It is to be hoped that repetitions of these experiments will overcome group performance differences of this kind by employing larger samples.

This replication of the original experiment helps to clarify some of the puzzling results obtained earlier. First, the incremental effect produced by the alien stimulus is seen more clearly. Secondly, the post-rest performance of all four groups of subjects conforms more closely to expectation in this second experiment. It remains for further studies to determine whether the unexpectedly poor post-rest performance of the experimental group in the first experiment reflects the operation of some unidentified variable or not.

It would appear from these results that the disinhibiting stimulus appears to produce its greatest effects when presented late in the practice period. This finding supports the general theoretical account of rotor performance in terms of the operations of reactive inhibition.

Further experiments on this topic should investigate in greater detail the temporal qualities of the disinhibitor, the varieties of disinhibiting stimuli and the effect of varied intensities of disinhibiting stimuli. It would also be of interest to discover whether the disinhibiting processes studied here operate in the same manner in psychological functions other than psychomotor behaviour. An unpublished experiment on visual perception suggests that this may not be the case.

### SUMMARY

A replication of the original experiment produced results similar in most respects to those obtained earlier and also helped to clarify some difficulties raised by the earlier results. In addition, it was found that the effect of the disinhibitor on performance and reminiscence is most marked when it is presented late in the practice period.

### REFERENCES

- AMMONS, R. B. (1955). Rotary pursuit apparatus: survey of variables. *Psychol. Bull.* **52**, 69-76.  
 BROADBENT, D. E. (1953). Noise, paced performance and vigilance tasks. *Brit. J. Psychol.* **44**, 295-303.  
 EYSENCK, H. J. (1956). Reminiscence, drive and personality theory. *J. Abnorm. Soc. Psychol.* **53**, 328-33.  
 EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.  
 EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory*. London: University of London Press.

- EYSENCK, H. J. & MAXWELL, A. E. (1961). Reminiscence as a function of drive. *Brit. J. Psychol.* **52**, 43-52.
- FRASER, D. C. (1953). Relationship of an environmental variable to performance on a prolonged visual task. *Quart. J. Exp. Psychol.* **5**, 31-2.
- HOVLAND, C. I. (1936). Inhibition of reinforcement and phenomena of experimental extinction. *Proc. Nat. Acad. Sci. Wash.* **22**, 430-3.
- JONES, H. G. (1960). Learning and abnormal behaviour. In Eysenck, H. J. (ed), *Handbook of Abnormal Psychology*. London: Pitmans.
- PAVLOV, I. P. (1927). *Conditioned Reflexes*. London: Oxford University Press.
- RAZRAN, G. (1939). Decremental and incremental effects of distracting stimuli upon the salivary CR. *J. Exp. Psychol.* **24**, 647-52.
- TREADWELL, E. (1960). The effect of depressant drugs on vigilance and psychomotor performance. In Eysenck, H. J. (ed.) *Experiments in Personality*, Vol. I. London: Routledge and Kegan Paul.

*(Manuscript received 28 April 1961)*





## IS THE SERIAL-POSITION CURVE INVARIANT?

BY ARTHUR R. JENSEN

*University of California*

The evidence presented by McCrary & Hunter and later researchers, which suggests that the form of the serial-position curve is constant throughout variations in a large number of factors that affect rate of learning, has been found inadequate to answer the question of the curve's invariance. Differences in percentage of total errors in serial learning data obtained under different experimental conditions would lead to the conclusion that the serial-position curve is *not* invariant under all the conditions for which invariance has been claimed.

Serial learning data may be analysed into three components—*difficulty* (number of trials to criterion), *efficiency* (percentage of errors), and *relative difficulty* of learning each position, i.e. the so-called serial-position effect. The main methodological inadequacy in studies of the serial-position effect is that the serial-position curve has always been confounded with either one or two of the other components, making it impossible properly to compare various serial curves based on data which differ either in *difficulty* or in *efficiency*. Plotting the serial curves simply in terms of the percentage of total errors at each position does not solve the problem; this method tends to eliminate the *difficulty* component but still leaves the curves confounded with the *efficiency* component.

An index of Relative Difficulty was proposed as a method of representing the serial-position effect. It has the advantages of not being confounded with the other components and of being the same shape whether it is based on errors or on correct responses. It results in what might be called a 'pure' serial-position curve and permits the direct comparison of serial-position curves obtained under various conditions. The Index is recommended as the only satisfactory method for comparing different serial-position curves. It is especially important in comparing serial-position effects in sets of data having different total error percentages. It is urged that future research on the serial-position effect adopt the proposed Index.

### I. INTRODUCTION

In scientific endeavour the discovery of a constant or an invariant function is always cause for joy. This must be especially true in psychology, where invariance is the rarest of findings. It is worthy of note, therefore, that in recent years the bow-shaped serial-position curve has seemed to be on its way to attaining the status of an invariant function. This interesting development deserves further looking into.

Since at least as far back as 1875, when Ebbinghaus began memorizing lists of nonsense syllables, it has been noted that the beginning and end of the serial list are learned most easily and the middle items are learned with greatest difficulty. If a list of items is learned to mastery and the number of errors made in the course of learning is plotted for each item in the series according to its position, the usual serial-position curve is obtained. It is almost always a skewed, bowed curve, with the maximum errors just past the middle of the list and with the last half of the list having more errors than the first half. This general phenomenon apparently emerges in serial rote-learning regardless of the nature of the items, which may be meaningful words, letters, nonsense syllables, proper names, objects, pictures, figures, etc.

A number of theories have been proposed to account for this so-called 'serial-position effect' (see McGeoch & Irion, 1952, pp. 125-34). According to various theories, the shape of the serial-position curve, particularly the degree of bowing, should be a function of such variables as distribution of practice, rate of presentation of the items, inter-item interval, degree of familiarity with the items, degree of similarity

among the items, and individual differences in learning ability. And, indeed, experiments on serial learning have shown that when the serial-position curve is plotted for groups of subjects in terms of the mean number of errors made at each serial position in the course of learning, the serial-position curve does appear to be a function of the above-mentioned variables.

Then McCrary & Hunter (1953) made an interesting discovery which seems to have become an important part of our general knowledge concerning serial learning (e.g. Woodworth & Schlosberg, 1954, p. 713). In essence, McCrary & Hunter found that when the serial-position curve is plotted, not in terms of the mean number of errors made at each position, but in terms of the percentage of the total errors that occurs at each position, none of the variables mentioned above had any effect on the shape of the serial-position curve. The McCrary & Hunter method of plotting the serial curve in terms of the percentage of total errors at each position, which thereby equates the area under all curves, seems reasonable if our interest is in the *shape* of the serial curve rather than in its absolute position on the ordinate. And if theories of serial learning make predictions about the relative difficulty of learning the items at different positions, it would seem reasonable that the errors at each position should be presented relative to the total number of errors that occurred. The percentage method of McCrary & Hunter accomplished this, and therefore should be more suitable than the method of mean errors for testing certain theoretical predictions concerning the *shape* of the serial curve. For example, Hull's theory of serial rote-learning (Hull *et al.* 1940) predicts that massed practice should produce a more bowed serial-position curve than distributed practice, because of a presumably greater accumulation of inhibitory potential in the middle of the list under massed practice. Hovland (1938) plotted serial-position curves for nonsense syllable learning under massed and distributed practice and under 2 and 4 sec. rates of syllable presentation. The curves, plotted as mean errors at each position, strikingly bear out Hull's prediction. But when McCrary & Hunter (1953) plotted these same curves on a percentage basis, the curves became practically identical. McCrary & Hunter showed that the manipulation of other independent variables also produced curves which differ greatly in shape when plotted as mean errors at each position but which assume almost identical shapes when plotted as percentage of errors. Braun & Heymann (1958) went further into this matter by performing an experiment which showed that the error curves (plotted on a percentage basis) for high- and low-meaningful lists of nonsense syllables did not differ. Nor did variations in inter-item interval or inter-trial interval have any effect on the shape of the curve in their study. Thus it would appear from these findings that we have possibly come upon our first 'constant' or 'invariant function' in the psychology of learning.

Indeed, such is the interpretation put upon this finding by some writers. On the basis of these data, for example, Murdock suggested that '...it would almost seem that the shape of the serial-position curve has nothing whatsoever to do with learning' (1960, p. 24) since it seems to be unaffected by variables that are known to affect rate of learning. The invariance of the serial-position curve is essential to Murdock's theory explaining the shape of the curve in terms of differences in the relative distinctiveness of the items in the serial list.

## II. A FALLACY

Actually, from the evidence in the literature, we are unable to determine whether or not the serial-position curve is invariant. The analyses of McCrary & Hunter (1953) and of Braun & Heymann (1958) are quite inadequate to answer this question. It is highly probable, however, from evidence in the Braun & Heymann study, that the serial-position curve is *not* invariant. That the idea of invariance even arose is mainly the result of a faulty analysis of what the serial-position curve is actually supposed to represent. It has apparently been assumed that there is just *one* serial-position curve for a given set of data, viz. the error curve, plotted either in terms of mean errors at each position or in terms of the percentage of total errors at each position. But the serial curve can also be expressed in terms of the number (or percentage) of *correct* responses at each position. It is sometimes plotted in this manner (e.g. McGeoch & Irion, 1952, p. 116). The important point, however, is that the *error* curve and the *correct* curve (whether in terms of absolute number or of percentage) for the same set of data are not always the same shape. It is possible for one curve to be much more bowed than the other. We cannot determine from the data in the McCrary & Hunter study how 'constant' the serial-position curves would appear if they had been plotted in terms of *correct* responses rather than in terms of *errors*.

Some contrived curves can illustrate the possible fallacy in the McCrary & Hunter analysis. Fig. 1 shows two fictitious but typical serial-position curves, *A* and *B*, plotted in terms of number of errors at each position. Curve *B* appears much more bowed than curve *A*. When these two curves are each expressed as the percentage of total errors at each position, they become *identical*, as shown in Fig. 2. This is what happened to all the curves in the McCrary & Hunter and the Braun & Heymann studies. But now look at the curves for *percentage correct* in Fig. 2. Plotted on this basis the curves are quite different in shape, curve *B* being much more bowed than curve *A*. The reason this happens is that curves *A* and *B* are each based on a different total error percentage. Curve *A* had 28.75% errors. Curve *B* had 57.50% errors.

If two or more percentage *error* curves are of the same shape but are based on different percentages or proportions of total errors (i.e. total errors/total responses) it is certain that the percentage *correct* curves are not of the same relative shapes. The percentage *error* curves in the Braun & Heymann study are all practically identical, but since they are based on quite different total error percentages, varying from 39 to 59%, it is impossible that their data, when plotted in terms of percentage correct at each position, could yield identical curves for the various experimental conditions they investigated. Unfortunately the percentage correct curves cannot be plotted with any exactitude from the data presented by Braun & Heymann. The point can be illustrated, however, by some of the writer's data on serial learning. From a group of sixty college students who had learned a 9-item serial list consisting of coloured geometric forms (triangles, squares, and circles coloured red, blue and yellow) the average serial-position curve of the ten subjects (group A) with the highest error percentage (mean = 57%) was compared with that of the ten subjects (group B) with the lowest error percentage (mean 33%). Fig. 3 shows the percentage error curves and the percentage correct curves for the two groups. In this case it is evident that the degree of bowing of the two curves is actually *reversed* for the two methods



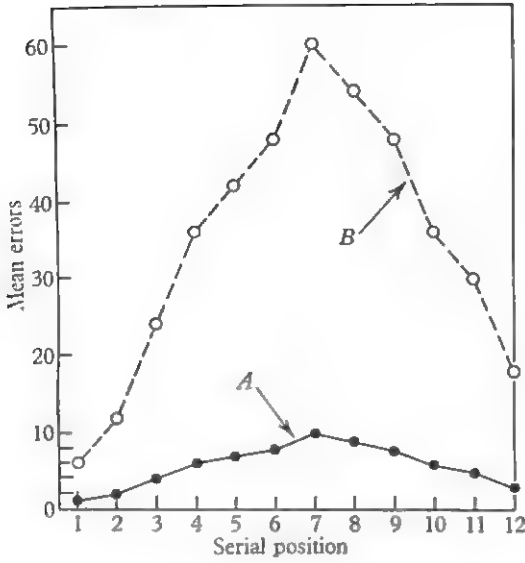


Fig. 1

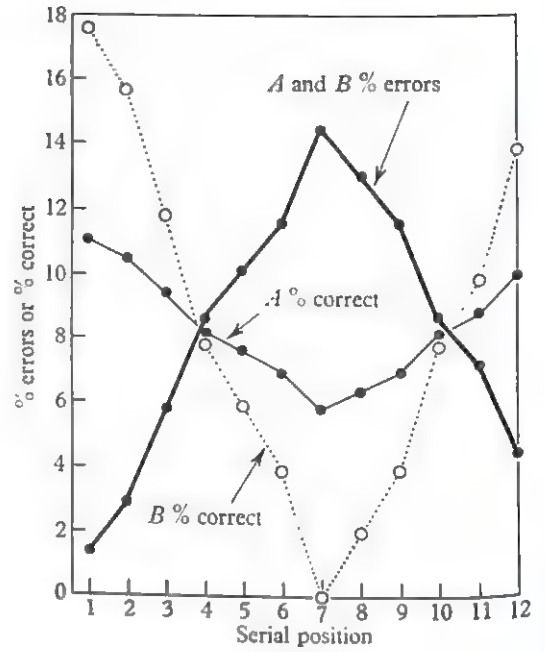


Fig. 2

Fig. 1. Two serial-position curves plotted in terms of number of errors at each position.

Fig. 2. The same 'data' as in Fig. 1, but here plotted in terms of the percentage of total errors at each position and the percentage of total correct at each position.

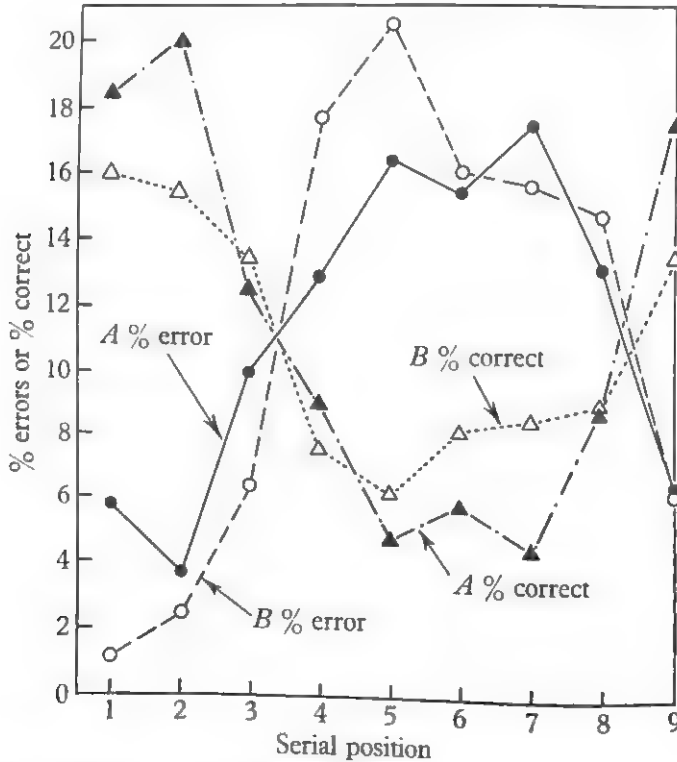


Fig. 3. Serial-position curves plotted in terms of percentage of errors and percentage correct at each position. Group A has a mean total error of 57 %, group B of 33 %.

of plotting curves. Curve *A* shows the least difference in shape for the two methods, because its total error percentage is closer to 50 % than that of group B. Only if both groups had the same error percentage would it be proper to compare their serial-position curves by only one method, i.e. either in terms of percentage of total errors or in terms of percentage of total correct at each position. To be more exact, one must insist that all subjects in each group have the same total error percentage if their averaged curves are to be properly compared. Without knowledge of these conditions, which most likely are not even approximated in the McCrary & Hunter and the Braun & Heymann studies, along with the absence of percentage correct curves, there is unfortunately no conclusion of any general importance to our understanding of the serial-position effect that can be drawn from these studies, other than the general fact that the serial-position effect is manifested under a variety of conditions.

### III. A SOLUTION

Serial rote-learning data may be analysed into three essential components:

- (1) The *difficulty* of the learning task, represented by the number of trials or stimulus presentations required to learn to a particular criterion.
- (2) The *efficiency* of learning, represented by the percentage of correct responses (or the percentage of errors) during the course of learning to a particular criterion.
- (3) The *relative difficulty* of learning the various positions. This actually is the serial-position curve.

Four methods of plotting serial-position curves are to be found in the literature, and all are inadequate in that they confound at least two or more of the above-mentioned components. These methods are applied either to errors or to correct responses, but error curves have been the most frequently used.

The methods are:

(a) Mean number of errors at each position. This curve confounds components 1, 2, and 3 above. It also has the disadvantage that subjects are weighted unequally in the means, so that the average curve for the group will tend more to reflect the shape of the curve for slow learners (i.e. those with many errors) than that of the fast learners. For testing theoretical predictions concerning the effects of different parameters on the shape of the serial curve, this method is unsatisfactory. However, it has been the most often used, especially in early studies.

(b) Percentage of errors at each position. This is determined by dividing the number of errors for all subjects at each position by the total number of errors for all positions. It has the effect of equating all curves for difficulty, i.e. it makes the area under all curves the same. But the shape of the curve is still confounded by the *efficiency* component. Furthermore, like curve *a*, subjects are weighted unequally in the group curve, those with a large number of errors determining the shape of the curve more than those with few errors. Thus, this method of representing the serial-position effect is also unsatisfactory. This was the method used by McCrary & Hunter (1953) and by Braun & Heymann (1958).

(c) Mean percentage of errors at each position. In this method the percentage of errors at each position for *each* subject is determined (by dividing the subject's total errors into the number of errors at each position) and these percentages are then

averaged to produce the group curve. This method is correct in weighting every subject equally in the average curve. And like curve *B*, it equates all curves on the difficulty component, making the areas under all curves the same. But the *efficiency* component is still confounded in the curve, which means that the error curve and correct curve will not necessarily have the same shape.

(*d*) Mean logarithm of errors at each position. This method has been used by Glanzer & Peters (1960). Since the work of McCrary & Hunter suggests that the total number of errors has a multiplicative effect on the shape of the serial-position curve, the transformation of the errors into log errors converts the multiplicative effect into an additive factor. This method has essentially the same effect as methods *b* and *c*. When it is applied to the McCrary & Hunter mean error curves, for example, it makes them all the same shape, as does the percentage method, but the curves still have different positions on the ordinate. This logarithmic method, however, is like the percentage method in that the efficiency component is still confounded in the serial curve and it is possible to obtain differently shaped curves for errors and for correct responses on the same set of data. In other words, in using methods *a*, *b*, *c*, or *d* a single curve cannot properly represent the serial-position effect. Comparing serial curves by analysis of variance, of course, does not overcome these shortcomings as long as errors, correct responses, or percentage errors are used in the analysis of variance.

Ideally, each of the three components of serial learning data should be kept separate from the others. What we really wish to know when we speak of the serial-position effect or the shape of the serial-position curve is the *relative difficulty* of learning each position, unconfounded by the absolute difficulty of the task (as measured by total trials) or the *efficiency* of learning (as measured by the total percentage of errors). Method *a* confounds all three variables, and methods *b*, *c* and *d* confound the *relative difficulty* and the *efficiency* components. If these components covaried perfectly, there would be no problem. But they do not. For example, in a group of sixty college students who had learned a 9-item serial list (coloured geometric forms) by the anticipation method (3 sec. rate of presentation, 6 sec. inter-trial interval) the following intercorrelations were obtained.

$r = 0.92$  ( $P < 0.01$ ) between *number of errors* and *number of trials* to attain the criterion of one perfect trial.

$r = 0.45$  ( $P < 0.01$ ) between *number of errors* and *percentage of errors*.

$r = 0.15$  (N.S.) between *number of trials* and *percentage of errors*.

There not only appear to be reliable individual differences in each of the components, but examination of the literature reveals that independent variables which affect the difficulty of the task (e.g. distribution of practice, rate of presentation, degree of intralist similarity, familiarity and meaningfulness of the items) also affect the efficiency of learning, the more difficult tasks having a higher error percentage.

### *The Index of Relative Difficulty*

In order that the serial-position curve shall represent what we actually mean by the serial-position effect, i.e. the *relative difficulty* of learning each position, unconfounded by the difficulty of the task or the efficiency of learning, an Index of Relative



Difficulty is here proposed as the only entirely satisfactory method of representing the serial-position effect. It has the advantage that it shows the *shape* of the serial curve independently of difficulty and efficiency. Also, the curves based on errors and on correct responses are always exactly the same shape, one being simply the upside-down mirror image of the other. This method permits valid direct comparisons of serial curves obtained under any conditions.

So that the serial curve of each subject is equally weighted in the average curve for the group, the Index is applied to the data for each subject. The points on the curve are then converted to percentages for each subject and are averaged over all subjects to produce the group curve.

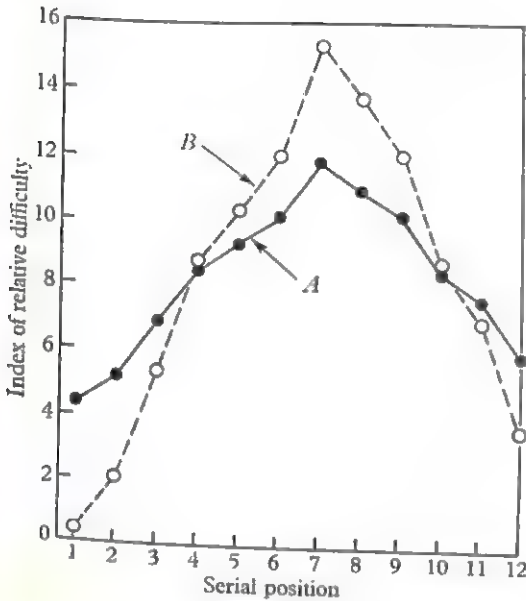


Fig. 4

Fig. 4. The data of Fig. 1 presented in the form of the Index of Relative Difficulty.

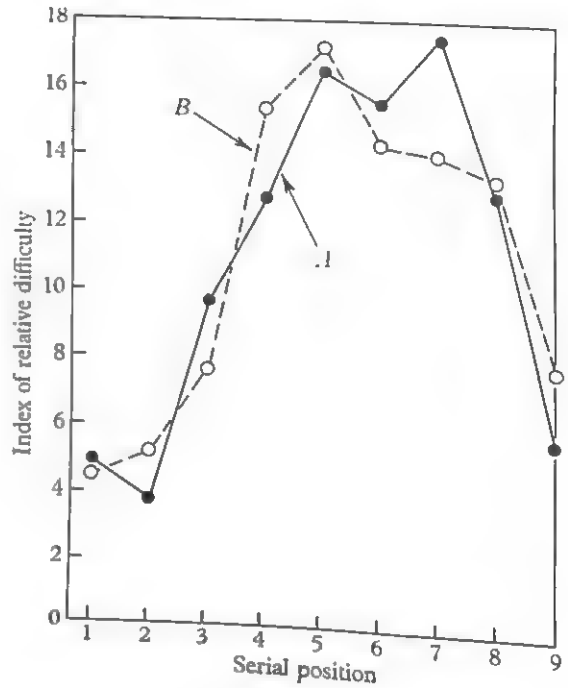


Fig. 5

Fig. 5. The data of Fig. 3 presented in the form of the Index of Relative Difficulty.

The method of obtaining the Index for a single subject is as follows:

(a) *Correction factor.* This, in effect, equates all curves for efficiency, i.e. total error percentage. When the error percentage is 50 %, the error curve and correct curve are of identical shape. Thus the method need only be applied to the errors.

$$\text{Correction factor}^* = \frac{\left( \frac{50}{\% \text{ errors}} \times \text{number of responses} \right) - \text{number of responses}}{\text{Number of positions in series}}$$

\* The % errors =  $100 \times \text{total number of errors} / \text{total responses}$ .

The number of responses = number of trials to criterion  $\times$  number of positions in series; in other words, the number of opportunities to make a correct response.

In actual computation the following formula, which is algebraically equivalent to that above, should be used:

$$\text{Correction factor} = 1/P(0.50 R - E),$$

where  $P$  = number of positions,  $R$  = total number of responses, and  $E$  = number of errors.

- (b) Algebraically add the Correction Factor to the number of errors at each position. (Note that the Correction Factor may be either positive or negative.)
- (c) After performing *b*, obtain the sum total over all positions.
- (d) Divide *c* into each position, *b*, to convert the quantities into percentages which will sum to 100%.

(e) The group curve is obtained by averaging *d* for each position over all subjects. The curve thus obtained is an Index of Relative Difficulty. Fig. 4 shows the same data as in Fig. 1 presented in the form of the Index. If the Index were computed from the correct responses rather than from the errors, the curves would simply be the upside-down mirror images of those in Fig. 4; in other words, they would contain identical information. We see that the two curves (*A* and *B*) in Fig. 4 are actually of different shapes and are *not* identical, as fallaciously represented by the percentage error curves in Fig. 2. From Fig. 5 we are able to make a valid statement concerning the relative degrees of bowing of the two curves, whereas on the basis of the percentage curves in Fig. 3 our statements about the relative degrees of bowing of curves *A* and *B* depend upon whether we are looking at the 'error' curves or the 'correct' curves. But for any one set of data there should be only one curve representing the serial-position effect, i.e. the relative difficulty of learning each position. This curve is provided by the Index of Relative Difficulty. When serial curves are compared by analysis of variance, the analysis should be performed on the Index scores.

#### REFERENCES

- BRAUN, H. W. & HEYMANN, S. P. (1958). Meaningfulness of material, distribution of practice, and serial-position curves. *J. Exp. Psychol.* **56**, 146-50.
- GLANZER, M. & PETERS, S. C. (1960). Re-examination of the serial-position effect. Mimeographed Technical Report, Research and Development Division, Office of the Surgeon General, Department of the Army.
- HOVLAND, C. I. (1938). Experimental studies in rote-learning theory. III. Distribution of practice with varying speeds of syllable presentation. *J. Exp. Psychol.* **23**, 172-90.
- HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T. & FITCH, F. B. (1940). *Mathematico-Deductive Theory of Rote Learning*. New Haven: Yale University Press.
- MCCRARY, J. W. & HUNTER, W. S. (1953). Serial position curves in verbal learning. *Science*, **117**, 131-4.
- MCGEOCH, J. A. & IRION, A. L. (1952). *The Psychology of Human Learning*. New York: Longmans, Green.
- MURDOCK, B. R. (1960). The distinctiveness of stimuli. (1960). *Psychol. Rev.* **67**, 16-31.
- WOODWORTH, R. S. & SCHLOSBERG, H. (1954). *Experimental psychology*. New York: Holt.

(Manuscript received 27 March 1961)

# SOME PROPERTIES OF BEHAVIOUR UNDER FIXED RATIO AND COUNTING SCHEDULES

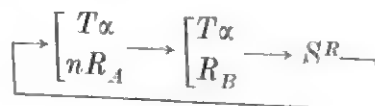
BY H. M. B. HURWITZ

*Birkbeck College*

The effects on rate of response and on a measure of the sequential properties of the response array of two reinforcement schedules, specified in terms of two response classes, lever pressing and tray responding, were investigated. Observations were made under conditioning and extinction. Under both conditions animals (rats) trained by the counting procedure responded at a consistently higher rate but received fewer reinforcements. The results are discussed with reference to two maze-running procedures and to double alternation studies.

Recent investigations have shown that systems of rewards and punishment, known as *schedules of reinforcement*, are major determinants of behaviour (see Ferster & Skinner, 1957; Sidman, 1960). Measures such as *rate of response* (Wilson, 1954; Kelleher, Fry & Cook, 1959), *topography of response* (Antonitis, 1951; Stanley & Aamodt, 1954; Hurwitz, 1954; Notterman, 1959; Schaeffer & Steinhorst, 1959; Millenson & Hurwitz, 1961), *sequential structure of behaviour* (Frick & Miller, 1951; Hurwitz, 1958; Mechner, 1958) and *behavioural stability* (Cummings & Schoenfeld, 1959; Hurwitz, Brener & Jones, 1961) all systematically respond to manipulations of the reinforcement schedule. Unlike many of the traditional independent variables investigated by psychologists, e.g. stimulus intensity, massing and spacing of practice trials, which can be specified in terms of single dimensional properties, schedules of reinforcement frequently involve complex contingencies. Mechner's (1959) recent efforts to provide a notation in terms of which such schedules and associated experimental operations can be specified is therefore to be welcomed as a contribution to the experimental analysis of behaviour. This notational system articulates similarities and differences in the specification of schedules more clearly than discursive descriptions are able to do.

The experiment reported here, which was completed before Mechner's notational system became known, investigated the effects of two schedules of reinforcement whose differences at first sight appeared trivial. Schedule FR specified that reinforcement—the presentation of food—would occur after a fixed number of responses of class A; schedule CS specified that reinforcement occurred only when a fixed number of responses of class A precede a response of class B. The schedules conform to the classical fixed ratio and counting type respectively. However, in both instances two responses of the behavioural flux terminating in reinforcement will be identified. In this way the results can be more effectively compared. The fixed ratio, FR, schedule is represented by

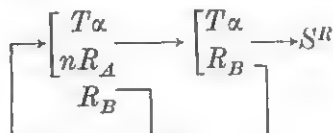


where  $R_A$  represents a response on the lever;  $R_B$  a response at the food tray;  $T$  a time period. The arrow connecting the bracketed events indicates that the responses



of the first bracket produce a second time event which sets the occasion for additional responses; the arrow connecting a response and a stimulus event, as in  $R_B \rightarrow S^R$ , indicates that a response has produced a stimulus event and the arrow connecting the reinforcer,  $S^R$ , to the beginning of the behavioural flow indicates a recycling operation as a whole or within segments of the schedule, as in schedule CS below.

The second schedule is the so-called counting procedure (Mechner, 1958). Reinforcement occurs only when a response  $R_B$  has been preceded by a fixed (or minimal) number of  $R_A$ 's. This procedure is represented as follows:



and involves an additional convention; the connecting link between  $R_B$  in the first bracket and the recycling arrow indicates that if  $R_B$  occurs after  $n-iR_A$ , a new series of  $nR_A$  is required to programme reinforcement. Under this schedule, therefore, different classes of responses have to be performed in a fixed order for reinforcement to occur.

The common feature of these two schedules is that reinforcement is dependent on the performance of a fixed number of responses of class A, lever pressing. The difference lies in the reset contingencies. The FR schedule resets on completion of the correct number of  $R_A$  responses, the CS schedule resets whenever a response of class B occurs.

In this study we were concerned to discover whether these differences in the schedules affected response frequency during reinforcement and experimental extinction sessions. We were also interested in what inferences could be drawn about the sequential structure of behaviour from the frequency data obtained. Rats were trained in a lever-pressing apparatus, food serving as reinforcement. The two response classes were lever pressing ( $R_L$ ) and contacts made with the food tray ( $R_T$ ).

## METHOD

*Subjects.* The subjects were ten experimentally naïve male hooded rats selected from the closed colony of animals maintained by the animal laboratory of the Department of Psychology. They were approximately 4 months old at the beginning of the experiment.

*Apparatus.* The apparatus consisted of a 12 in. square box, a food pellet dispenser which delivered 0.04 g. pellets into a recessed food tray, and a lever mounted 2½ in. above floor level and 5 in. to the right of the food tray. The lever, when depressed with a downward force of 5 g., closed a micro-switch. The food tray was shielded by a transparent door which had to be nuzzled open to gain access to food. The circuits were arranged to deliver the pellet when the tray was opened provided the required number of lever responses had previously occurred. A buzzer sounded on every lever press.

The apparatus was permanently housed in a large sound-insulated chamber.

The controlling electronic circuits could be programmed in either of two ways: (a) the FR programme made food available whenever the tray response ( $R_T$ ) followed three lever responses ( $R_L$ ); (b) the CS programme contained a reset arrangement such that the counting device set to zero whenever a tray response occurred before the required number of lever responses had been emitted.

Magnetic counters tabulated lever responses, tray contacts and reinforcement frequencies.

*Habituation and feeding schedule.* Three weeks before training the animals were placed into an automatic feeding battery which made food available once daily for 2 hr.

The subjects all received the same initial training in the apparatus. They were first given four preliminary training sessions. On each session five reinforcements only were given for performing the tray responses. On the first session three pellets were placed in the tray. A further five pellets were delivered by the experimenter operating an external switch. The presentation of each pellet was preceded by an auditory stimulus—a 0.5 sec. buzz.

On the fifth day each subject was placed in the apparatus for half an hour. The circuits were programmed to deliver a pellet whenever a lever response preceded a tray response. If the subject delayed between lever and tray for 20 sec., the reinforcement automatically reset. The initial training was terminated after the subject had secured no less than 70 pellets within two successive sessions.

The second phase of the experiment lasted for five half-hour sessions. The subjects were divided into two groups, group CS consisting of four subjects, group FR of six. The reinforcement schedule for group CS was arranged so that when the lever responses preceding a tray response fell short of the required number, namely three, the programme reset. Reinforcement was thus dependent on the correct sequence of the two response classes. For group FR no such restriction applied. In practice, then, under FR the subjects could shuttle back and forth between lever and tray, yet be certain of reinforcement on the third shuttle.

The final phase lasted for 2 hr. sessions and involved an extinction procedure. During these sessions the food dispenser was emptied of pellets.

### RESULTS

The rate of lever and tray response during reinforcement sessions and extinction sessions, the rate of lever responses to number of reinforcements, and the total number of reinforcements over the five sessions of the second phase of the experiment were analysed for each group. Table 1 summarizes the rate of response data.

Table 2 represents the ratio  $R_L/S^R$ , i.e. the mean number of lever responses per reinforcement during the final session of phase 2. The results show the two groups

Table 1. Mean rate of lever and tray response ( $R_L$  and  $R_T$  respectively) during reinforcement and extinction sessions under a counting (CS) and fixed ratio (FR) reinforcement schedule of training

	Reinforcement sessions		Extinction sessions	
	$R_L$	$R_T$	$R_L$	$R_T$
Group CS	191.1	151.0	344.5	152.2
Group FR	116.4	146.3	189.0	242.5
Significance of difference	$P < 0.01$	N.S.	$P < 0.01$	$P < 0.02$

Table 2. The ratio of lever responses ( $R_L$ ) to total number of reinforcements ( $S^R$ ) on the final training session under a counting and fixed ratio schedule of reinforcement for each subject

Counting schedule		Fixed ratio schedule	
Subject	$R_L/S^R$	Subject	$R_L/S^R$
21	8.05	31	3.09
22	5.43	32	3.26
63	5.43	33	3.22
64	4.33	34	3.33
		35	3.17
		39	3.17

to be different with respect to this variable,  $P = 0.02$ . Finally, the total number of reinforcements secured during the counting schedule and fixed ratio schedule phase of the experiment were compared and failed to show that the groups differed on this measure.

#### DISCUSSION

The analysis of food-reinforced lever pressing by rats may be viewed in terms of the conditioning of a response sequence. Such a sequence involves many components such as moving towards the lever, pressing and releasing it, moving towards the food tray, retrieving food, etc. The result of consistent positive reinforcement following the emission of such a heterogeneous response array is twofold: it changes the probability that component responses of the sequence will be repeated under similar stimulus conditions, i.e. the rate of response is increased; it may affect the sequential order of response. This latter idea finds expression in Hull's (1932, 1943) theorizing under the heading of the 'goal gradient hypothesis'. The notion that units of behaviour larger than individual responses become reinforced in a free operant situation is, of course, an old one. Skinner (1938) discussed it in his early publications, and it has since been put forward by several writers under a variety of labels, e.g. operant chaining, response-unit formation, sequential pattern learning, etc. (e.g. Mowrer & Jones, 1943; Keller & Schoenfeld, 1950; Ratner, 1956; Mechner, 1958). Yet there are few experimental studies which deal directly with this phenomenon by making observations of changes in the sequential order of response emission (Frick & Miller, 1951; Mechner, 1958; Millenson & Hurwitz, 1961). In general, sequential re-organization has been inferred from single-response frequency data. In this study however, the frequency measure has been extended to two responses of the presumptive chain. The additional information gained and summarized above should enable us to relate data and inference more closely than is often the case.

Consider first the counting schedule. Under this schedule, reinforcement occurs only when responses are emitted in a predetermined order. In this respect the schedule is like a non-correction maze-running procedure in which a trial is terminated whenever either a correct or an incorrect choice is made. The measure of learning is expressed as the ratio of correct to the total number of choices made. A similar measure can be defined in the lever-pressing apparatus. Given the total number of lever responses and the number of reinforcements, the ratio  $R_L/S^R$  indicates that, on the average, 65% of the response sequences occurred in the required (correct) order of three lever responses followed by one tray response (Table 2).

Secondly, under the counting procedure, consistently more lever than tray responses were performed; the reverse being true under fixed ratio. Differential effects of the conditioning of lever and tray responses are also reflected in the extinction scores. Thus if the ratios of lever to lever plus tray responses are considered, group CS maintained a higher ratio during both reinforcement and extinction sessions than group FR. This suggests that the reinforcement procedures acted not only differentially on the probability of response measure but also enforced different sequential arrangement of the responses. In a separate experiment, where a similar fixed ratio schedule was under investigation and direct measures of sequential structure were taken, this prediction was substantiated. There it was found that most



animals shuttled back and forth between lever and tray in a simple alternation pattern after as many as 1000 reinforcements (Hurwitz, 1962). However, when the schedule was changed so that an exteroceptive, response-produced stimulus marked the end of  $nR_L$ , the response pattern underwent a systematic change and took the form of behaviour observed under a counting schedule.

Thus it appears that the two schedules, counting and fixed ratio, affect the emission of tray responses in different ways. Under a counting schedule, tray responses appear to be under the control of discriminative stimuli which are different from those operative under fixed ratio. Since suitable exteroceptive stimuli have been deliberately excluded, the control of tray responding—at least for the counting schedule group of animals—must have passed to movement-produced stimuli. Once again this analysis makes contact with maze-learning studies, specifically with studies on the question of central and peripheral factors in the control of maze performance (see Munn, 1950). In many of these experiments it was concluded that whilst the pattern of left-right choices could not be acquired by rats in the absence of exteroceptive stimuli which signal each change in the response requirement, once the maze has been mastered the maze-habit could be sustained by movement-produced stimuli only. (Racoon, cats, rhesus monkeys and chimpanzees do very much better than rats in mastering such problems entirely without exteroceptive stimuli.) At the time, these findings were interpreted as showing that response-produced stimuli can take over discriminative functions from exteroceptively provided signals.

In the majority of these maze studies a correction procedure was used. Thus, provided the correct number of choices was made and irrespective of the number of incorrect choices, reinforcement took place. This procedure is analogous to the fixed ratio procedure and the results of the two types of experiment appear to be highly comparable.

But even though early maze-running experiments throw doubt on the discriminative self-sufficiency of movement-produced stimuli, recent investigations of behaviour under a variety of reinforcement schedules have provided overwhelming evidence in support of the opposite conclusion. One thinks of the typical behaviours which emerge under fixed interval schedules, counting schedules, mixed schedules and differential rate of response type schedules (e.g. Ferster & Skinner, 1957). The problem which therefore needs to be raised and investigated concerns the isolation of the conditions which enable movement-produced stimuli to gain functional control over the emission of responses. This is not only a taxonomic problem but a parametric one. One needs to determine quantitatively the effect of different conditions on the slope of the discrimination learning curve. Two studies in which it was shown that movement-produced stimuli gained discriminative control over an operant response under a fixed ratio type schedule illustrate this point. Schlosberg & Katz (1943) used a double alternation schedule in a lever-pressing apparatus. A single lever was used which had to be pushed twice in one direction and then twice in the opposite direction. At first simple alternation patterns were observed but in time most subjects adopted double alternation, although reinforcement would also occur whenever the correct number of responses had been totalled. Similarly, in a study by the present writer (Hurwitz, 1962) it was found that *some* rats developed counting behaviour under fixed ratio conditions, although their performance was generally inferior. Thus,

the reinforcement contingencies appeared to determine the rate of conditioning discriminations based on movement-produced stimuli. They do not determine this phenomenon in all-or-none fashion. Schlosberg & Katz ran their animals daily for 4 months and reported that only towards the tail-end of this period was double alternation observed. Hurwitz ran his animals for fourteen half-hour sessions and again observed changes from simple alternation between lever and tray to counting-like behaviour only towards the final stages of the experiment.

In general, the task we set ourselves in this study—to discover the behavioural effects of two closely allied schedules of reinforcement—has met with success. There appear to be pronounced differences between the behaviours initially generated under a fixed ratio and a counting schedule, although in time and under special stimulus conditions such differences may disappear. Further, the two schedules produced different absolute and relative rates of response for each of two response classes, and this also occurred under a subsequent experimental extinction procedure. By inference this suggests that the schedules resulted in different sequential response arrangements. This latter variable has as yet received insufficient attention within operant conditioning methodology and remains to be systematically related to many of the isolated independent variables whose effects, at least on the rate of a single response, are already well established.

#### REFERENCES

- ANTONITIS, J. J. (1951). Response variability in the white rat during conditioning, extinction and reconditioning. *J. Exp. Psychol.* **42**, 273–81.
- CUMMINGS, W. W. & SCHOENFELD, W. N. (1959). Some data on behaviour reversibility in a steady state experiment. *J. Exp. Anal. Behavior*, **2**, 87–90.
- FERSTER, C. & SKINNER, B. F. (1957). *Schedules of Reinforcement*. New York: Appleton-Century.
- FRICK, F. C. & MILLER, G. A. (1951). A statistical description of operant conditioning. *Amer. J. Psychol.* **64**, 20–36.
- HULL, C. L. (1932). The goal gradient hypothesis and maze learning. *Psychol. Rev.* **39**, 25–43.
- HULL, C. L. (1943). *Principles of Behavior*. New York: Appleton-Century.
- HURWITZ, H. M. B. (1954). Response duration of lever pressing in the rat. *Quart. J. Exp. Psychol.* **6**, 62–71.
- HURWITZ, H. M. B. (1955). Response elimination without performance. *Quart. J. Exp. Psychol.* **7**, 1–7.
- HURWITZ, H. M. B. (1958). A source of error in estimating the number of reinforcements in a *Bull. Brit. Psychol. Soc.* **43**, 2A.
- HURWITZ, H. M. B. (1962). The effect of a discriminative stimulus on behaviour under a lever-pressing apparatus. *J. Exp. Anal. Behavior*, **2**, 149–52.
- HURWITZ, H. M. B., BRENER, J. & JONES, B. (1961). Some experiments on behavioural stability. fixed ratio schedule of reinforcement. (In preparation).
- KELLEHER, R. T., FRY, W. & COOK, K. (1959). Inter-response time distribution as a function of differential reinforcement of temporally spaced responses. *J. Exp. Anal. Behavior*, **2**, 91–106.
- KELLER, F. S. & SCHOENFELD, W. N. (1950). *Principles of Psychology*. New York: Appleton-Century.
- MECHNER, F. M. (1958). Probability relations within response sequences under ration reinforcement. *J. Exp. Anal. Behavior*, **1**, 109–21.
- MECHNER, F. M. (1959). A notation system for the description of behavioral procedures. *J. Exp. Anal. Behavior*, **2**, 133–50.
- MILLENSON, J. R. & HURWITZ, H. M. B. (1961). Some temporal and sequential properties of behavior during conditioning and extinction. *J. Exp. Anal. Behavior*, **4**, 97–106.

- MOWRER, O. H. & JONES, J. (1943). Extinction and behavior variability as a function of the effortfulness of task. *J. Exp. Psychol.* **33**, 369-86.
- MUNN, N. L. (1950). *Handbook of Psychological Research on the Rat*. New York: Houghton Mifflin Co.
- NOTTERMAN, J. M. (1959). Force emission during bar-pressing. *J. Exp. Psychol.* **5**, 341-7.
- RATNER, S. C. (1956). Effect of extinction of dipper-approaching on subsequent extinction of bar-pressing and dipper-approaching. *J. Comp. Physiol. Psychol.* **49**, 576-81.
- SCHAEFFER, H. H. & STEINHORST, R. A. (1959). The effect of changing the schedule of reinforcement upon duration of responding. *J. Exp. Anal. Behavior*, **2**, 335-41.
- SCHLOSBERG, H. & KATZ, A. (1943). Double alternation lever pressing in the white rat. *Amer. J. Psychol.* **56**, 274-82.
- SIDMAN, M. (1960). *Tactics of Scientific Research*. New York: Basic Books.
- SKINNER, B. F. (1938). *The Behavior of Organisms*. New York: Appleton-Century.
- STANLEY, W. C. & AAMODT, M. J. (1954). Force of responding during extinction as a function of force requirement during conditioning. *J. Comp. Physiol. Psychol.* **47**, 462-4.
- WILSON, M. P. (1954). Periodic reinforcement: interval and number of periodic reinforcements as parameters of response strength. *J. Comp. Physiol. Psychol.* **47**, 51-56.

(Manuscript received 1 April 1961)





## AN EXPERIMENTAL STUDY OF THE GROWTH OF SOME LOGICAL STRUCTURES

BY K. LOVELL, B. MITCHELL AND I. R. EVERETT

*University of Leeds*

A number of experiments of the type suggested by Piaget & Inhelder's book *La genèse des structures logiques élémentaires* were given to a population of primary and educationally sub-normal (E.S.N.) special school children. The findings among the former group generally confirm those of the Genova school. It has been possible to extend Piaget & Inhelder's findings by giving a number of tests to the same children, and by making a comparative study of primary and E.S.N. school pupils.

### I. INTRODUCTION

The ability to classify has long been recognized as an important aspect of the cognitive processes. From the time of Binet questions involving classification have been found in intelligence tests; Goldstein & Scheerer (1945) used tests of classification to study the lack of flexibility of thinking said to occur amongst certain clinical cases; more recently Semeonoff & Trist (1958) have given details of a number of tests involving classification together with the mode of administration and assessment. Lovell (1955) using group versions of the Wisconsin, Trist-Hargreaves and Vinacke Sorting Tests with a non-clinical population of adolescents and young adults, showed that there is an ability involved in classifying the above-named and other materials, over and above the well recognized *g*, *verbal* and *k* abilities. In the work of Piaget & Inhelder (1959), however, there is a major attempt to analyse the successive forms and stages through which children pass in classifying material, and to relate the development of classification to the growth of logical structures.

Piaget himself in his earlier work studied the problem of the logical subset in a verbal context, but in his later research he became more convinced that classifications and seriations were the result of mental operations (i.e. actions which have become internalized, which are carried out in the mind and are reversible). He considered that the formation of concepts and the underlying systems of classification was due to something more fundamental than socio-linguistic communication, language being a necessary but not a sufficient condition for the formation of classes and the structure of relationships. The work of Oléron (1956) supports this view to some extent, for he found that deaf mute children completed elementary operations of classification, but were retarded in more complex conceptual systems. Luria & Yudovich (1959) also report that an improvement in language and verbal formulation assisted the classifying process in Twin A ('Yura').

This paper describes an experimental study involving the individual testing of children, using experiments of the type described in Piaget & Inhelder's book *La genèse des structures logiques élémentaires*, together with three experiments which bear a close resemblance to certain of those given in Piaget's book *The Child's Conception of Number*. Only the essentials of the study can be given here, details of

the presentation of materials, instructions, further questions asked, and other data obtained, may be obtained from the authors. It is assumed that access can be had to Piaget & Inhelder's book.

## II. PROCEDURE, RESULTS AND DISCUSSION

### 1. *First series of experiments*

In this part of the study five experimental techniques were used. The population consisted of 10 children, representative of all levels of ability, in each of the age groups 5 to 10 in a primary school, and 10 children in each of the age groups, 9, 11, 13, 15 in F.S.N. (Educationally Sub-normal) special schools. It is not possible to say how many tests were given to any one group of children in the Geneva studies.

#### *Experiment 1. Spontaneous classifications of geometrical shapes and letters: additive classification*

The material consisted of five groups of shapes, each containing six elements. The groups were composed of squares, triangles, circles, rings and half rings. All six elements in any one group were identical in shape and size. Two elements were made of blue and one of pink plastic foam; three elements were made of manilla of which two were covered with red, and one with blue plastic tape. The material also included six letters of the alphabet. The letters L, T, P were cut in plastic foam, the first in pink, the last two in blue; while the letters L, X, T were cut in manilla covered with plastic tape, the first two in red and the last in blue. All the elements were presented in disorder in front of the subject and the general instructions were: 'Put together things that go together' or an essential variation of this. Later, subjects might be told to put the elements in 'a better order' or to arrange the elements in 'a better way'.

Stage 1 classifications gave the elements arranged in short or long lines, or in small incomplete groups. The latter were extremely unstable and the subjects often placed elements together, took them away, or rearranged the spatial arrangement of the elements after they had declared 'I've finished'. Two squares might be placed together in a short line, then a circle and a square. A long line might alternate between shape and shape, colour and colour, material and material. Collections were determined more by spatial shape than by inclusive elements; Piaget termed this type of classification 'figural'. At Stage 2 (non-figural collections) classifications showed a breaking away from spatial arrangement and there was more emphasis on similarities and differences among the elements. It is in effect a stage between figural collections and classes representing inclusion. Classification proper was found at Stage 3.

Each of the three stages can be divided into sub-stages although it is our intention here only to state the subdivisions for Stage 3. At Stage 3a all the elements are classed together according to one criterion; namely, shape, material, colour, form. They are then subdivided according to a further criterion. Stage 3b indicates that all the elements are classified and subdivided according to two criteria, and at Stage 3c all the elements are classified and subdivided according to three criteria. The number of children at Stage 3c should be comparable with the numbers of



children making three classifications in Experiment 6; this point will be referred to again when that experiment is discussed.

The results for Experiments 1, 2, 3 and 4 are combined in Table 2.

### *Experiment 2. Multiplicative classification*

The material consisted of sixteen cards, 2 by 2 in., on which pictures of rabbits were painted; eight cards showed similar rabbits running, 4 being painted black and four being painted white; eight cards showed similar rabbits sitting, four being painted black and four painted white. There was also a box 10 by 6½ in. painted black inside, and a similar-sized one painted white inside, and a further box 9 by 12 in. subdivided into four equal sections by movable partitions.

The sixteen cards were presented in disorder and the following kinds of instructions were given, and questions asked:

- (1) 'Put together those that are alike, those that go together.'
- (2) Present the black and white boxes. 'Put some in this box and some in that box.'
- (3) Present the box with the four partitions. 'Put together those that are alike and put them in different parts of this box.'
- (4) One partition of the box used in (3) was taken out, subdividing the rabbits into black running and sitting in one half, and white running and sitting in the other half. 'Can you put these (indicating the black running rabbits), with these (indicating the black sitting rabbits)?' 'Why?' 'Can you put these with these (indicating the white running and white sitting)?' 'Why?'
- (5) The first partition was replaced and second partition removed, subdividing the rabbits into black running and white running in one half, and black sitting and white sitting in the other half. 'Can you put these (indicating the black running rabbits) with these (indicating the white running rabbits)?' 'Why?' 'Can you put these with these (indicating the black sitting and white sitting)?' 'Why?'

In (4) and (5) the rabbits are classified in dichotomies; by colour, i.e. white and black; then by shape, i.e. running and sitting. The questions determine whether the subject understands that once the four subclasses have been separated they can be reunited; in (4) by additive classes black running, black sitting, also white running, white sitting; in (5) by additive classes black running, white running, also black sitting, white sitting. If the subject is able to place a subclass in both dichotomies, he has shown the ability to understand multiplicative classification.

Using the above approach we observed the subjects' spontaneous arrangements and by questioning we could determine how the latter had been made.

At Stage 1 we find obvious figural arrangements, i.e. mixed groups of black and white, sitting and running rabbits. We also found arrangements in four subclasses in accordance with perceptual similarities but without any evidence of the extension of a unified class. Thus although the subject may relate black running and white running rabbits in perceptual relationships, when asked 'Can we put these with these?', he might respond 'No'. At Stage 2 there is some extension of a class, but there is no evidence of a multiplicative system either in spontaneous arrangement in boxes, or in response to questions. Finally at Stage 3, subjects consistently show relationships of a multiplicative nature. Thus we find arrangements in the box such as

White running		Black running
White sitting	+	Black sitting

and questions are answered with adequate proof.

*Experiment 3. Anticipation and visual seriation*

The subjects were first presented with four cardboard squares all of different sizes, and they were asked to 'put the squares in order'. Any child who was unsuccessful was helped, and when the series was constructed he was questioned further to ascertain if he understood what 'putting in order' meant.

When it was clear that the child had understood his instructions, he was presented with ten rods in a certain disordered arrangement, the same for each child. The rods were of 1 sq. cm. cross section, their lengths ranged from 1 to 10 cm.; and they were of ten different colours (they were in fact, the Cuisenaire materials). A base line was provided on a piece of cardboard.

The experimenter said to the child, 'Look at these sticks carefully. You are going to put them in order. Show me which one will go here, which will go next and so on.' The child had first to anticipate the series (without moving the sticks) that he later had to construct, and the task of the experimenter was to trace the development of the coordination of anticipation and past actions.

At Stage 1 figural type solutions were found. Both the anticipations and actions indicated a lack of coordination of intension and extension in relation to the rods. For example, a child G.H. age 5 years 2 months anticipated the rods in the sequence (of length) 6, 1, 4, 2, 9, 10, 3, 5, 8, 7. He then made the following series: 8, 2, 9, 10, 5, 4, 16, 7, 3. At Stage 2 we find anticipations and constructions partly correct, while at Stage 3, pupils both anticipated and constructed the series correctly.

*Experiment 4. Multiplication of asymmetrical transitive relations*

The material consisted of sixteen cardboard squares each  $2\frac{1}{2}$  by  $2\frac{1}{2}$  in. Each square contained a painting of a tree leaf. The leaves were in four different sizes and four different shades of green so that they could be arranged simultaneously according to size and shade. If I, II, III, IV indicate order of size, and 1, 2, 3, 4 indicate order of colour, the elements were arranged as in Table 1. In addition there were duplicate cards for II 2, III 2, I 1, II 4.

Table 1. *Arrangement of elements in the table illustrating Multiplication of Asymmetric Transitive Relations*

	1	2	3	4
I	I 1	I 2	I 3	I 4
II	II 1	II 2	II 3	II 4
III	III 1	III 2	III 3	III 4
IV	IV 1	IV 2	IV 3	IV 4

I, II, III, IV, indicate order of size; 1, 2, 3, 4, indicate order of colour.

The sixteen cards were presented to the subject in disorder and he was asked to arrange them as he thought best. A series of standard instructions and questions was used according to the nature of his spontaneous reactions. The instructions and questions included:

'Please put them in order.' 'Can you put them in a better way?' 'Can you make more groups?' 'Can you put the dark ones together and the light ones together?'

'Can you put them so that you can find the big leaves right away?' 'Can you put them so that you can find the same colour right away?'

If the subject was unable to construct the table, the elements of one dimension were arranged for him and he was asked to put the others in order. If he was unable to do this a second dimension was constructed for him and the child was again asked to complete the table.

When the subject had completed the table, element II2 was withdrawn and the four duplicate elements were presented with the instruction, 'Put in the one which will go here'.

At Stage 1 there was an absence of seriation, and the intension and extension of classes were not coordinated. During Stage 2 there was evidence of differentiation and partial coordination of intension and extension. At Stage 3 the operation of multiplicative asymmetrical transitive relations was complete.

Table 2. Number of children at Stages 1, 2, 3 in experiments: Addition of Classes, Multiplication of Classes, Visual Seriation, Multiplication of Asymmetrical Transitive Relations

Age in years	No. of child- ren	Stage 1				Stage 2				Stage 3			
		A	M	VS	MTR	A	M	VS	MTR	A	M	VS	MTR
Primary school													
5	10	3	5	1	3	7	5	6	5	0	0	3	2
6	10	1	0	1	0	6	8	5	7	3	2	4	3
7	10	0	0	0	0	6	5	3	7	4	5	7	3
8	10	0	0	0	0	0	0	1	1	10	10	9	9
9	10	0	0	0	0	2	2	0	1	8	8	10	9
10	10	0	0	0	0	1	1	2	1	9	9	8	9
E.S.N. school													
9	10	3	3	7	6	7	7	2	4	0	0	1	0
11	10	0	0	2	0	7	6	6	8	3	4	2	2
13	10	0	0	0	0	4	5	2	4	6	5	8	6
15	10	0	0	0	0	1	3	0	6	9	7	10	4

A, Addition of Classes; M, Multiplication of Classes; VS, Visual Seriation; MTR, Multiplication of Asymmetrical Transitive Relations.

#### Experiment 5. The hierarchical classification of animals

The materials consisted of three toy ducks; three toy birds that were non-ducks (robin, turkey, hen); and five toy animals that were non-birds (pig, dog, sheep, goat, horse). There were also box A (3 by 3 in.); box B (6 by 6 in.); and box C (9 by 9 in.); each of these had 1 in. high transparent sides. The boxes were placed one inside the other so that when the animals were placed in the respective boxes, inclusive relationships could be observed. Thus animals (non-birds) go in box C, birds (non-ducks) in box B, and ducks in box A.

Only a selection of the questions asked is given below. These relate to general questions of inclusion, (a) to (f); quantification of inclusion, (g) to (j); further quantification of inclusion, (k) to (n):

(a) 'Can you put the ducks in this box (B)?' 'What is this box (B) for?'

(b) 'Can you put the ducks in this box (C)?' 'What is this box (C) for?'



- (c) 'Can you put these birds (indicating birds non-ducks) into this box (C)?'
- (d) 'Is it right to put these (indicating non-birds) into this box (B)?'
- (e) 'Is it right to put these (indicating birds non-ducks) into this box (A)?'
- (f) 'Is it right to put these (indicating non-birds) into this box (A)?'
- (g) 'All these (indicating all the animals) belong to a farmer. Has he more ducks than birds, or less?'
- (h) 'Has he more birds than animals, or less?'
- (i) 'Are there more birds or more animals?'
- (j) 'Are there more ducks or more birds?'
- (k) 'If you killed all the birds will there be any animals left?' 'Which?'
- (l) 'If you killed all the ducks will there be any birds left?' 'Which?'
- (m) 'If you killed all the animals will there be any birds left?' 'Which?'
- (n) 'If you killed all the birds will there be any ducks left?' 'Which?'

The results of the first series of experiments are now discussed under two heads.

(1) Table 2 shows that, with the apparatus used in these experiments, the ability to achieve Stage 3—that of operational mobility—is achieved in addition of classes, multiplication of classes, visual seriation and multiplicative asymmetrical transitive

Table 3. *Numbers of children responding correctly to questions of class inclusion and hierarchical classification involving animals*

		Question number and type of question					
Age in years	No. of children	(a) All ducks birds?	(e) All birds ducks?	(b) All ducks animals?	(c) All birds animals?	(f) All animals ducks?	(d) All animals birds?
Primary school							
7	10	9	8	2	2	9	9
8	10	10	8	3	2	10	10
9	10	10	9	1	1	10	10
10	10	10	10	5	5	10	10
E.S.N. school							
9	10	3	6	3	3	6	6
11	10	6	7	3	4	10	9
13	10	6	6	1	4	10	9
15	10	7	7	5	7	10	10

relationships at about the same time in primary school children. This is the first time that this has been shown through giving the four tests to the same children. Although Piaget & Inhelder say that the ability to perform addition and multiplication of classes appears *pari passu*, they do not present data from the same children to support their claim. The increase in operational mobility between 7 and 8 years of age also confirms the view of the Geneva school.

In all four tests, children at Stage 1 showed no consistency in their actions. Indeed, there seemed to be a complete lack of ability to see and plan the next step. At Stage 2, however, the beginning of a construction seemed to be remembered, for there was some coherence between one move and the next. But there was no sudden jump to Stage 3 as one might expect if visual perception was relied upon exclusively. Rather, it seemed that the move from Stage 1 to Stage 3 was closely related to the

improvement in the coordination of retroactions and anticipations, attention being increasingly directed to and from the partial solutions. Because of the nature of the materials and procedures it was not possible to estimate the influence of language. From our observations it would appear that language plays only a small part in the formation of these types of logical structures, but we cannot be sure since language

Table 4. *Numbers of children responding correctly to questions involving quantification in relation to inclusion*

Age in years	No. of children	Question number and type of question						
		(g) More ducks than birds?	(h) More birds than animals?	(g) + (h) correct	(i) More birds or more animals?	(j) More ducks or more birds?	(i) + (j) correct	All questions correct
Primary school								
7	10	0	0	0	0	0	0	0
8	10	3	3	3	1	2	1	0
9	10	3	4	3	0	0	0	0
10	10	6	6	6	2	3	2	1
E.S.N. school								
9	10	0	0	0	1	1	1	0
11	10	0	0	0	0	0	0	0
13	10	1	0	0	0	0	0	0
15	10	2	2	2	1	2	1	0

Table 5. *Numbers of children responding correctly to questions of quantification in relation to inclusion*

Age in years	No. of children	Question number and type of question						All questions correct
		(n) Kill all birds any ducks left?	(l) Kill all ducks any birds left?	(n) + (l) correct	(m) Kill all animals any birds left?	(k) Kill all birds any animals left?	(m) + (k) correct	
Primary school								
7	10	4	9	4	2	8	2	1
8	10	8	10	8	2	8	2	2
9	10	10	10	10	1	10	1	1
10	10	9	10	9	5	10	5	5
E.S.N. school								
9	10	4	8	3	4	4	1	0
11	10	1	8	1	2	5	0	0
13	10	5	10	5	5	8	4	4
15	10	5	10	5	4	9	3	4

might have been used covertly. However, we are inclined to agree with the Geneva school that language is not in itself a sufficient condition for the formation of these structures; for if classification depended greatly on language, logical addition should be easier than logical multiplication since sentence structure, e.g. 'The black and the white rabbits ran' should tend to aid logical addition.

The performance of 15-year-old E.S.N. special school pupils is about equal to that of 7- to 8-year-old children in primary schools. This is frequently found in the other experiments reported here and clearly parallels the findings of Lovell & Slater (1960) in respect of the development of the concept of time in E.S.N. special school children. It is noticeable that visual seriation was rather easier, and multiplicative asymmetrical transitive relations rather more difficult for the 15-year-old E.S.N. children than the other experiments. The former finding is probably due to the fact that these children have had some years of this type of activity both within and without school.

(2) Table 3 shows the numbers of children giving correct responses to questions of inclusion of the type, 'Are all the ducks birds?', and 'Are all the animals birds?'. Many subjects answered correctly with a 'yes' or 'no' but were unable to provide a satisfactory reason; these were not assessed as passing the item. The hard questions are those of the kind 'Are all the ducks animals?' and 'Are all the birds animals?' when these questions are related to the action of putting ducks or birds in the class of animals containing non-birds as well. It seems as if birds, including ducks, are not included within the class of animals for many primary school children. Piaget & Inhelder (1959) certainly asked 'Are all the birds animals?' ('Les oiseaux sont des animaux?'), and we kept close to the form of questioning laid down by the Geneva school. It was only afterwards that we became wise to the difficulty the children experienced. At the beginning of the experiments when all the animals were together each child was asked 'What are all these?' or similar. The usual reply was 'Animals' or 'farm animals'. But when the groups were separated, many subjects no longer looked upon birds as animals. Perhaps the Geneva children were abler, or language or slight culture pattern differences affect the issue. The E.S.N. children at 15 are again equal to the 7- to 8-year-old primary school children in performance except that they do much better in questions (b) and (c). Perhaps sheer experience of life has helped these older E.S.N. pupils to appreciate that ducks are animals.

In Table 4 one can see how difficult questions of quantification of inclusion are when cast in certain forms. As soon as these questions were posed the children seemed to ignore the concrete materials in front of them and instead to deal with the problem at an abstract verbal level. This is confirmed in our view by the figures in Table 5 where the concreteness of the situation was maintained to some extent by asking what the results would be if certain classes of animals were killed. In these instances the numbers of correct responses rose.

Once again there is an increase in the number of correct responses between 7 and 8. Questions (b) and (c) in Table 3 and question (m) in Table 5 again reveal the difficulty children have in looking upon birds as animals. It is very doubtful if the small number of correct responses to these questions is due to lack of operational mobility *per se*.

This experiment involving animals is not a satisfactory one although at first sight it would appear to provide great interest. It is believed that the class of animals is more abstract than most of the other classes involved in this study. Perceptual relationships are not evident, and these do have some relevance since concrete logical operations cannot be dissociated from the intuitive content of the elements to which the operations apply. The early unstable and indefinite verbal concept does not help the children to appreciate the significance of their actions in relation to animals.



(pig, dog, robin, duck, etc.) which they do not usually think of as being associated in real life. They are familiar with squares and circles being associated, and with different kinds of flowers found in a bunch of cut flowers or growing in a garden, but not at first with a rather odd collection of animals. Thus retroactive and anticipatory processes have no significance in relation to the class of animals as a whole in this experiment.

## 2. *Second series of experiments*

In this part of the study five experimental techniques were used. The population consisted of 10 children, representative of all levels of ability, in each of the age groups 5, 6, 7, 8, 9-11 years in a primary school, and 25 children in each of the age groups 10-13 and 14-15 years in E.S.N. special schools. The primary school was situated in a somewhat lower socio-cultural area than was the primary school used in the first series of experiments. In retrospect we have reason to believe that the 8-year-old primary group was rather less able than the other year groups.

### *Experiment 6. Visual classification*

The material consisted of four large blue squares of side 5 cm.; four small blue squares of side  $2\frac{1}{2}$  cm.; three large blue circles of diameter 5 cm.; three small blue squares of diameter  $2\frac{1}{2}$  cm., and a large red circle of 5 cm. diameter. Later, the following material was added if necessary; one large red square of side 5 cm.; one small red circle of  $2\frac{1}{2}$  cm. diameter; one small red square of side  $2\frac{1}{2}$  cm.

The procedure followed these general lines:

(a) The first group of elements was placed in front of the child and he was asked to put them into groups as he liked. This was to give him preliminary practice in sorting the elements.

(b) The child was asked to put the elements into two groups, e.g. 'Can you put these in two lots that are different from each other'. The child was further questioned to find out his criterion. The subjects usually grouped by shape.

(c) He was then asked to find another way of sorting the elements into two groups. The subjects often grouped by size on this occasion.

(d) The child was asked to find a third way of arranging the elements in two groups. This was often the unique class, which consisted of all the blue pieces in one group, and the one red piece forming the other group. The subject had, of course, to attribute to the one red square the properties of a class. If he was unable to make the unique class, the second group of elements was added and he was again asked to make the third classification. No child tried to make the unique class for his first classification; four made it for their second classification but none of these was able to make a third classification.

The results for Experiments 6 and 7 are displayed in Table 6, the figures in brackets indicating the numbers of children who could sort the material into two groups in three different ways without the addition of the three red pieces; i.e. these figures give the numbers of children who could form the unique class as one of three ways of classifying the material. These numbers, when added to the corresponding numbers under the other columns headed V, give a total of 10 for each primary school age group.

*Experiment 7. Tactile-kinaesthetic classification*

The material consisted of sixteen elements all made of wood. These comprised two balls of 4 cm. diameter; two balls of 2 cm. diameter; two cubes of side 4 cm.; two cubes of side 2 cm.; two circles of 4 cm. diameter; two circles of 2 cm. diameter; two squares of side 4 cm.; two squares of side 2 cm. All the squares and the circular discs were 0.2 cm. in thickness. A frame was made so that a cloth could be draped over it in such a way that the child could not see objects placed behind the cloth, but the experimenter could observe the actions made by the child in handling the objects. Furthermore, the top of the frame was low enough for the child and interviewer to see one another; conversation was thus helped.

Table 6. *Numbers of children who were able to make 0, 1, 2 or 3 classifications, by visual and tactile-kinaesthetic perception*

Age in years	No. of children	No classifications		One classification		Two classifications		Three classifications	
		V	TK	V	TK	V	TK	V	TK
Primary school									
5	10	2	2	6	3	1	4	1(1)	1
6	10	0	0	4	5	6	4	2(0)	1
7	10	0	0	4	1	4	6	4(2)	3
8	10	1	0	4	5	3	2	4(2)	3
9-11	10	0	0	0	0	6	4	6(4)	6
E.S.N. school									
10-13	25	1	1	14	16	9	4	1(1)	3
14-15	25	1	1	15	10	7	8	2(2)	6

V, visual perception; TK, tactile-kinaesthetic perception.

The objects were placed in a tray behind the cloth and the child told to feel the objects; to remember their 'feel', and to place them in one or other of two boxes (also behind the cloth) that he had seen earlier. Thus:

(a) The child was asked to place the objects in the two boxes so that there would be the same sort in each box.

(b) The objects were placed in the tray again, well mixed, and the child told to find another way of sorting the objects into the two boxes.

(c) The objects were placed in the tray a third time, well mixed, and the child was asked to find a third way of sorting the objects into the boxes.

The three classifications made were shape ('rounds' v. 'squares'); size ('big' v. 'little'); volume or third dimension ('thick' v. 'thin').

*Experiment 8 (a). Additive composition of classes involving marked perceptual differences*

The materials consisted of twenty plastic 'poppet' beads, of which eighteen were brown and two white. The beads were laid out in disorder on a piece of blue paper. The child was instructed to pick up some of the beads and examine them; it was pointed out at the same time that the beads were of the same material throughout and not

just painted on the surface. A series of questions were put to the subject and a selection of these is given:

'Are there more brown ones or white ones?'

'Are there more plastic beads or are there more brown ones?'

'If I made a necklace of white beads, and I made another necklace of brown beads, which necklace would be the longer?'

'If I made a necklace of the plastic beads, or if I made a necklace of brown beads, which necklace would be the longer?'

In the case of some children who did not answer the last question correctly, they were asked to make the actual necklaces and the question was repeated. In no case did the child change his answer.

Table 7. *Numbers of children giving correct responses to questions involving additive compositions of classes and relations of inclusion using coloured beads and flowers*

Age in years	No. of children	Stage 1		Stage 2		Stage 3	
		Beads	Flowers	Beads	Flowers	Beads	Flowers
Primary school							
5	10	10	10	0	0	0	0
6	10	5	6	2	2	3	2
7	10	4	7	2	1	4	2
8	10	4	4	3	3	3	3
9-11	10	0	1	0	0	10	9
E.S.N. school							
10-13	24	14	17	6	4	5	4
14-15	25	9	14	7	7	9	4

#### Experiment 8 (b)

Experiment 8 (a) was repeated with materials which were well known and which had specific names. The children were presented with twenty red roses and three daffodils, all the flowers being artificial. The names of the flowers were clearly stated. If a particular child called the daffodils, say, daisies, this was accepted and the name kept throughout. Great care was taken to establish that the terms 'red' and 'yellow' were understood. A series of questions were put to the child, a selection of which is given:

'If they were growing in a garden like this (flowers standing upright) and you wanted a very big bunch, must you pick the flowers or the roses?'

'If I pick the daffodils what will be left?'

'Are there more flowers or are there more roses?'

At Stage 1 subjects were unable to grasp the relationship between the part and the whole as well as between the parts themselves. Stage 2 was a transitional stage in that they sometimes answered correctly and sometimes did not; while at Stage 3 they were able to relate the parts to one another and to the whole in each question.

#### Experiment 9. *The use of the words 'all' and 'some' in situations involving colour and weight*

A box totally enclosed a lever balance, the latter controlling the positions of an apple concealed within the box. The apple appeared only when the 'heavy' boxes



weighing 5 oz. each were placed on the platform, the remaining 'light' boxes weighed  $4\frac{1}{2}$  oz. each. There were three light red boxes, three light blue boxes and three heavy red boxes, all the boxes being of the same size. The following procedure was followed:

(a) The experimenter said, 'Some of the boxes placed here (platform) make the apple appear, some boxes do not' (demonstrated).

(b) 'Place each of the boxes here (indicating platform) one at a time'.

(c) 'Place the boxes here (indicated) that make the apple appear, and place the boxes there (indicated) that do not make the apple appear.' The subjects had to place the boxes side by side on their respective parts of the table and not on top of one another.

Four preliminary questions were then asked to focus the pupils' attention on to certain relevant variables:

(i) 'Why did you place the boxes here?' (indicating those that made the apple appear).

Table 8. *Numbers of children responding correctly to questions involving the use of the words 'all' and 'some' in situations involving colour and weight*

Age in years	No. of children	<i>Hr</i> All heavy ones red?	<i>bl</i> All blue ones light?	<i>rH</i> All red ones heavy?	<i>lb</i> All light ones blue?	<i>Hr</i> + <i>bl</i>	<i>rH</i> + <i>lb</i>	All correct	<i>Hr</i> + <i>lb</i> (W)
Primary school									
5	10	4	9	8	5	4	4	2	2
6	10	7	10	10	7	7	7	6	6
7	10	9	8	10	7	7	7	5	7
8	10	9	6	8	8	5	6	4	7
9-11	10	8	10	10	10	9	10	8	8
F.S.N. school									
10-13	25	18	17	20	13	11	10	3	7
14-15	25	15	17	18	17	8	17	7	12

(ii) 'Why did the apple appear?' (Here the child has to abstract the concept of weight.)

(iii) 'Did the red boxes make the apple appear?'

(iv) 'Did the blue boxes make the apple appear?'

The main questions followed. These included:

'Are all the heavy boxes red?' (*Hr*) 'Why?'

'Are all the blue boxes light?' (*bl*) 'Why?'

'Are all the red boxes heavy?' (*rH*) 'Why?'

'Are all the light boxes blue?' (*lb*) 'Why?'

Each question was asked twice and the second of the two answers was recorded. By this means children were able to 'correct' their first answer if they could in much the same way as the children in Geneva did. The results are displayed in Table 8.

The results are now very briefly discussed under a number of heads.

(1) Table 6 shows that among the primary school population, the numbers of children making three classifications is about the same whether visual or tactile-

kinaesthetic perception is used. When the unique class is one of the three classifications, made visually, tactile-kinaesthetic perception gives rather better results. This is generally in keeping with Piaget's prediction although we have demonstrated the fact by giving both tests to the same group of children. Among E.S.N. special school children three classifications were made rather more easily using tactile-kinaesthetic perception compared with visual perception. We have clear evidence that all children tend to 'search' less among the variables using visual perception.

Some children could make classifications but were unable to define the criterion verbally. The E.S.N. children especially had difficulty in naming the common quality of the elements. This again suggests that simple classificatory processes do not greatly depend on language, although later a verbal term to describe a class may well allow the child to extend his operations.

The number of children making three classifications in these two experiments is comparable, as one would expect, with the number of children at Stage 3c in Experiment 1 where the elements were classed and subdivided according to three criteria. The number of primary school children in each age group who were at Stage 3c of Experiment 1 is given below:

Age in years	5	6	7	8	9	10
Number of children	0	0	0	4	3	6

The corresponding figures for E.S.N. special school children:

Age in years	9	11	13	15
Number of children	0	0	0	3

These experiments involving classification make it clear that the degree of operational mobility available to the child is reflected in his ability to co-ordinate the increasing number of variables when criteria are changed.

(2) It appears from Table 7 that the attainment of Stage 3 in the additive composition of classes and relations of inclusions is rather more difficult when using the flowers than when using the beads. As the flowers were brightly coloured there was greater perceptual contrast than with the beads, and this might be responsible. The number of 8 year olds reaching Stage 3 is small; this might be due to the fact that 8 year olds in our population were rather a poor group. On the other hand, Hyde (1959) reports only 7 out of 48 European children at Stage 3 in an identical experiment using beads. Judged against her figures, our 8 year olds did well.

(3) Table 8 reveals that it is rather easier to deal with the questions of the type 'Are all the blue (colour) ones light?' than 'Are all the light (weight) ones blue?' When reference to abstract weight comes first in the question there is a tendency to reverse the question. As was explained earlier, the subjects were asked each question twice and so were allowed to 'correct' their first answers just as Piaget & Inhelder permitted their subjects to do. The infant and the E.S.N. children frequently changed a right answer to a wrong one, whereas the older junior children corrected wrong answers to right ones. Even so, answers to individual questions tend to be unstable and a better indication of the understanding of the pupils is obtained from columns 5 to 8 reading from left to right.

## III. CONCLUSION

Over-all the results reported here agree fairly well with those of Piaget & Inhelder. It has been possible to confirm many of their predictions by giving a number of tests to the same pupils: these pupils being drawn from a known population of school children. In addition the work has been extended to cover E.S.N. special school pupils and it has shown the limited ability of the pupils to develop logical structures.

We acknowledge the great co-operation that we received from the head teachers of the schools that were involved in the study. Our thanks are also due to the children who patiently tried to answer the many questions that they were asked.

## REFERENCES

- GOLDSTEIN, K. & SCHEERER, M. (1945). *Tests of Abstract and Concrete Thinking*. New York: Psychological Corporation.
- HYDE, D. M. (1959). An investigation of Piaget's theories of the development of the concept of number. Unpublished Thesis, University of London Library.
- LOVELL, K. (1955). A study of the problem of intellectual deterioration in adolescents and young adults. *Brit. J. Psychol.* **46**, 199-210.
- LOVELL, K. & SLATER, A. (1960). The growth of the concept of time: a comparative study. *Journal Child Psychology and Psychiatry*, **1**, 179-90.
- LURIA, A. R. & YUDOVICH, F. I. A. (1959). *Speech and the Development of Mental Processes in the Child*. London: Staples.
- OLÉRON, P. (1956). *Recherches sur le Developpement Mental des Sourds-Muets*. Paris: C.N.R.S.
- PIAGET, J. & INHELDER, B. (1959). *La Genèse des Structures Logiques Élémentaires*. Neuchatel: Delachaux and Niestle.
- SEMEONOFF, B. & TRIST, E. (1958). *Diagnostic Performance Tests: A Manual for Use with Adults*. London: Tavistock.

(Manuscript received 8 May 1961)



## EFFECTS OF A SUBSIDIARY TASK ON PERFORMANCE INVOLVING IMMEDIATE MEMORY BY YOUNGER AND OLDER MEN

By D. E. BROADBENT

*Applied Psychology Research Unit, Medical Research Council, Cambridge*

AND ALASTAIR HERON

*Occupational Aspects of Ageing Research Unit, Medical Research Council, Liverpool*

Many practical situations expect a man to seek out and respond to signals, while at the same time remembering which signal is the one now requiring response. Most situations devised by psychologists, on the other hand, examine either reaction to signals without memory, or memory without reaction. A task has been devised in which both functions are involved. It consists of sets of random numbers which are observed through a small slot allowing only a few numbers at a time to be visible. The subject has to cross out particular digits, and has to remember which digit he is seeking at any particular instant. This type of task is compared with a more conventional number crossing task in which the subject has to find all instances of one particular digit, and both types of task are performed with and without an additional auditory distraction. In the present experiment, conducted with groups of young and of old subjects, it was found that the tasks involving memory are very vulnerable to distraction. The older subjects differed greatly among themselves, some being seriously inferior in performance to younger subjects; when no such difference in variance appears in the similar task without memory load. When a distracting task is presented together with a main task involving memory, older subjects do very badly at one or the other, whereas young subjects do reasonably at both.

### I. INTRODUCTION

It is a commonplace of everyday life that one is called upon to look for 'signals' of various kinds while remembering which particular signal is the important one at any given moment. Sometimes this remembering involves information one has only just acquired: at other times, it is more long-term in character. Further, the searching task may often have to be carried out in far from ideal conditions; and it is sometimes the case that what are loosely termed 'distractions' actually contain a significant source of information which must be taken in at the same time as one is trying to use already-stored information.

Modern evidence about immediate memory suggests that it is a separate process from the long-term sort, and very vulnerable. This view was of course suggested by cyberneticists, because of the existence of separate types of storage in computers, one type of store holding information which was needed relatively permanently, while another type holds information needed only for short periods. Positive evidence for a similar distinction in man began to be accumulated by Brown (1958) who showed that information held in short-term memory was relatively little affected by lapse of time between presentation and recall, but badly affected by reaction to other stimuli requiring action between presentation and recall. There was an interaction between the amount being held in memory and the amount of intervening activity: if a lot is being held in memory, very little interpolated material is needed to disturb recall, but if only a little is being held, quite a reasonable amount of intervening activity can take place.

These results suggest a view of immediate memory as involving continuous action on the part of some mechanism which is also needed for dealing with perception, and which cannot do both at once. One might think of a trace which fades away as fast as it is established, and requires to be strengthened at intervals by rehearsal: or one might think of the stored information as passing round a recurrent circuit which periodically runs through a channel used also in the perception of fresh information. Such interpretations are supported by results such as those of Broadbent (1957). In this case six digits were presented to one ear, and two others to the other ear, the listener being instructed to repeat first the six and then the two. If the two were simultaneous with the last pair of the six, performance was much better than if they arrived at an earlier time. In the latter case, of course, the trace would have time to decay before rehearsal could take place, and a time of 1 or 2 sec. was apparently enough to cause such a deterioration, in marked contrast to the effects of time on memory when rehearsal is possible.

A similar demonstration was that of Conrad (1957) who presented material for immediate recall at various speeds and also required the recall to be at set speeds. Over the range of speeds where no rehearsal is possible between items, recall is inferior when the speed of presentation and recall is lower. In the latter case more time is allowed for the trace to decay between rehearsals. For this reason a dial telephone is inferior to a set of numbered keys, as a way of reproducing numbers (Conrad, 1958).

Kay (1953) showed that if a continuous series of stimuli is presented, one after the other, and a man is asked to respond not to the stimulus now present but to the one before last, great difficulty arises. This might from some points of view be surprising, as the man is merely being asked to remember two items as well as the present one, but the difficulty is intelligible if one remembers that any rehearsal of the stored items will be incompatible with perceiving the present one.

Such investigations indicate the peculiar character of short-term memory: but in addition there is more than a suspicion that this function is badly affected by increasing age. Kay (1953) himself, for instance, showed that this effect was greater in people over thirty than in younger persons. Conrad's experiment has been repeated by Fraser (1958) on both young and old Canadian school-teachers; while the group were very similar in immediate memory at a fast rate, the older subjects did much worse when the rate was slow and when fading of the trace was presumably therefore more important. Lastly, Inglis (1960) has presented digits simultaneously to the two ears of old patients with memory disorders, as well as those without such symptoms. While the two groups did equally well on the digits on the ear recalled first, the patients with memory disorders did much worse than the controls did on the information on the other ear, which perhaps had longer to decay before rehearsal.

In this paper we report the results obtained by using simple tasks designed to require the combined use by the subject of memory and of direct perception. The experiment was conducted first with groups of young subjects, and then repeated with older groups, so that the effects of age might be examined.

## II. SUBJECTS

(a) *Younger subjects*

Six groups of young naval ratings were formed: their ages ranged from 18–25 years. Details of age, intelligence, emotional stability and sociability appear in Table 1, from which it can be seen that while the groups were fairly well matched for age and the personality variables, random allocation did not succeed so well where intelligence was concerned. The numbers in certain groups are also rather small.

(b) *Older subjects*

Five groups, of 8 men each, were selected from the 45–60 age-group of the M.R.C. voluntary subject panel at Liverpool. All these 40 men were in Social Class III of the Registrar-General's classification. The five groups were selectively equated for mean age (51 years) and vocabulary level: it will be seen in Table 1 that the use of this double criterion resulted in further close similarities between these experimental groups when non-verbal intelligence, emotional stability, and sociability are considered.

Table 1. *Subjects, and allocation of tasks*

Type of task	...	N	ND	S	SD	L	LD
Younger subjects							
N		7	6	7	6	4	4
Age	Mean	20.00	20.17	20.43	20.83	19.63	19.25
	$\sigma$	2.78	1.86	1.18	2.54	2.17	1.64
AH4	Mean	72.43	73.17	70.86	60.00	83.00	76.00
	$\sigma$	16.57	12.07	13.30	15.20	10.08	9.72
H <sub>1</sub>	Mean	8.86	8.17	7.29	9.17	7.13	4.25
	$\sigma$	5.05	2.67	3.49	2.97	3.37	2.17
H <sub>2</sub>	Mean	4.57	4.00	5.43	4.83	3.63	4.75
	$\sigma$	1.92	2.24	2.72	2.34	1.22	3.90
Older subjects							
N		8	—	8	8	8	8
Age	Mean	51.88	—	51.75	51.50	51.25	51.63
	$\sigma$	2.26	—	2.73	3.57	4.20	4.33
Vocab.	Mean	24.00	—	24.25	24.00	24.63	24.50
	$\sigma$	5.29	—	4.38	4.21	4.27	4.12
PM	Mean	43.88	—	42.63	42.63	43.13	42.38
	$\sigma$	6.37	—	8.76	7.41	7.00	4.00
H <sub>1</sub>	Mean	6.13	—	8.50	8.75	6.38	5.63
	$\sigma$	3.80	—	2.83	3.03	3.20	2.06
H <sub>2</sub>	Mean	5.25	—	5.36	6.13	6.38	5.63
	$\sigma$	3.46	—	3.46	1.90	2.73	2.50

Details of experimental populations. Key: AH4 and PM (Progressive Matrices, 1938), tests of non-verbal intelligence; Vocab., Mill Hill, Senior B synonyms; H<sub>1</sub>, H<sub>2</sub>, scores for emotional instability and unsociability from Heron's two-part personality inventory.

## III. THE EXPERIMENTAL TASKS

Three tasks were prepared: one involving short-term memory, one involving long-term memory, and a third which did not involve memory at all. The latter was for comparison purposes, and was meant to provide a check on the importance of the memory factor. Since the other tasks were derived from this 'no memory' one, it will be simplest to describe that one first and then to indicate the new features introduced into the other tasks.



*(a) The 'no memory' task*

The person being tested was given a pile of foolscap sheets of paper, each having on it a random series of numbers. The numbers were typed in ten-digit rows, one row beneath another, and each page contained two columns each consisting of thirty such rows. Thus there were 600 digits per page. The man was also provided with a sheet of cardboard containing a slot which was exactly the size of one of the ten-digit rows. Thus with the cardboard over the paper he could only see one row of numbers at a time.

He was told that his job was to put a pencil line through every digit of a particular type: say, through every figure 2. He was to start with the top row of the left-hand column on the first sheet of paper, work his way down, then work down the right-hand column, then repeat the process for the next page, and so on till all the numbers had been examined. One practice page was given; after that, nine more were done as the main test in a continuous run. The experimenter noted the time taken for each page separately.

*(b) The short-term memory task*

In this task the same material was provided, but the man was told that while he was to start by crossing out, say 2's, he would find periodically that he would come to some other figure which had a red circle round it. When this happened, he was to stop crossing out whatever figure he had been seeking previously, and start looking instead for numbers like the one with the red circle round it. In due course he would come to yet another red circle, and then he was to start crossing out whatever number was in that circle, and so on. Thus the man might start by crossing out 2's, go on to 6's, then to 8's, then 5's and so on: and at any stage he had to remember which number he was trying to find. Owing to the slot, he could not see the last red circle once he had passed it. Circles were in fact inserted every ten rows on the average, so that over fifty changes occurred in the nine pages of the main test. As before, a practice page was given: this was particularly important in this task in order to make sure that the instructions were understood.

*(c) The long-term memory task*

The task we have just mentioned does indeed involve memory, but only of a very temporary kind. The difficulty of devising a task involving long-term memory is that the experimenter can hardly practise people on an artificial job to the high level often expected in real life. This difficulty was met by making use of a memory which everybody has thoroughly acquired in ordinary life: memory for the sequence of numbers from 1 to 10.

The same sheet of paper and cardboard were used as before, but this time the man was told to start by crossing out the first figure 1 he saw, then the next figure 2, then the next figure 3, and so on. When he reached 9 he looked for 0, then for 1 again, and so on. Thus in addition to looking for a number he had again to remember what to look for: but in this case the memory of the sequence ought to be better established.

One final point should be made about this task: every other figure 1 of those which the man should have crossed out was marked with a red circle, and he was told to restart his sequence at that point if he had somehow got out of step with the correct sequence. This was done to simplify scoring: for this task as for the other two tasks, correctly marked sheets were used for comparing with each man's results, and discrepancies noted. But if no steps had been taken to get him back in step after any error in the long-term memory task, all subsequent crossings-out would have differed from those on the answer sheet.

*(d) The distracting task*

This task consisted of a tape recording in which a letter of the alphabet was said every 5 sec. In every ten such letters, one occurred twice: after each ten a voice asked the subject to write down the letter that had been repeated. A 10 sec. gap was then left for him to do this, which he did at the side of his answer-sheet for the main task. This task was combined, or not, with each main task, giving a 'no memory' task (N), a 'no memory plus distraction' (ND), a 'short memory' (S), and so on.

## IV. RESULTS\*

## (a) Types of error

(i) *Isolated Omissions (IO)*. In all three tasks there were of course occasional isolated errors of omission, where the man had failed to cross out a number he ought to have (Table 2). There were also errors of commission, where a single incorrect number had been crossed out, but these were most exceptional.

Errors of this type were not significantly increased by distraction on either the short-term memory task (S/SD) or on the long-term memory task (L/LD), in either younger or older subjects.

(ii) *Red circles not seen (RCNS) and seen but forgotten (RCF)*. Both the tasks involving memory, however, showed other types of error which did not appear in the 'no memory' task. In the short-term memory task, there were a number of cases where a man missed a whole succession of numbers and at the same time crossed out a succession of incorrect ones. These episodes always ended at a red circle: and are most easily explained as being due to the man either failing to see a circle: and are most easily explained as being due to the man either failing to see a red circle or else seeing it but later forgetting what number he was supposed to be checking. In some cases it could be shown that the man had in fact made some correct responses after passing a particular red circle, and only then had entered

Table 2. *Isolated omissions*

	N			S			SD			L			LD		
	N	Mean	$\sigma$	N	Mean	$\sigma$	N	Mean	$\sigma$	N	Mean	$\sigma$	N	Mean	$\sigma$
Younger	7	9.1	4.8	7	16.3	15.7	6	25.5	14.3	4	8.1	5.4	4	15.5	9.6
Older	8	11.3	7.5	8	9.1	4.6	8	11.0	5.2	8	11.0	7.4	8	19.8	13.5

into a series of errors. These cases must definitely be due to forgetting, whereas cases in which the sequence of errors began at a red circle might possibly be due to failure to see the circle. The incidence of both types of error is shown in Table 3. (Since this task involves memory, it should perhaps be noted that similar results on young subjects have also been found with an auditory task that did not involve memory. The interference between the distraction task and the main task does not depend upon the necessity for memory in both of them.)

Sequence errors of these two kinds were significantly increased by distraction (S/SD) in both younger and older subjects.

(iii) *Forgetting resulting in omission (F/O) or in repetition (F/R)*. In the long-term memory task, there appeared a type of error in which a number was crossed out twice running or else was not crossed out at all in its proper place in the sequence. All these errors must have been due to forgetting, since there can be no question that the subjects knew the correct sequence of numbers when they started the test. Table 3 gives the incidence of each type of error.

Sequence errors of these two kinds were significantly increased by distraction (L/LD) in both younger and older subjects.

\* The test of statistical significance used is that of Whitfield, J. W., 'Rank correlation between two variables, one of which is ranked, the other dichotomous' (*Biometrika*, 1947, 34, 292-6) and observed differences are described as significant when  $P < 0.05$ .

(b) *Speed of work*

It will be seen from Table 4 that:

(i) No significant difference in speed was observed between the no-memory and the short-term memory task without distraction (N/S) in either younger or older subjects.

(ii) Distraction slowed performance on the short-term memory task (S/SD) among both younger and older subjects (even though, as will be seen, most older subjects failed to carry out the distraction task).

Table 3. *Sequence errors*

Error		S			SD		
		N	Mean	$\sigma$	N	Mean	$\sigma$
Red circles not seen	Y	7	0.7	1.4	6	3.3	3.3
	O	8	0.5	0.7	8	1.1	1.1
Red circles forgotten	Y	7	0.7	1.0	6	3.8	4.0
	O	8	0.6	0.9	8	2.5	2.2
		L			LD		
		N	Mean	$\sigma$	N	Mean	$\sigma$
Forgetting/omission	Y	4	0.5	0.5	4	5.8	3.6
	O	8	1.1	1.6	8	6.1	4.0
Forgetting/repetition	Y	4	0.9	0.8	4	3.8	2.9
	O	8	1.5	1.5	8	3.4	2.7

Table 4. *Time (sec.)*

	N			S			SD			L			LD		
	N	Mean	$\sigma$	N	Mean	$\sigma$	N	Mean	$\sigma$	N	Mean	$\sigma$	N	Mean	$\sigma$
Younger	7	17.2	3.6	7	16.7	1.4	6	22.1	2.1	4	20.1	0.5	4	25.9	1.6
Older	8	20.9	1.9	8	18.5	3.2	8	27.1	6.6	8	28.2	6.6	8	26.5	6.8

(iii) Time taken for the long-term memory task was significantly longer than for short-term memory (L/S) among both younger and older subjects.

(iv) Distraction slowed performance on the long-term memory task (L/LD) among the younger subjects only (here again most older subjects failed to do the distraction task).

(v) Within-group variability (differences between individuals) increases from no-memory to short-term memory without distraction (N/S) and again for the addition of distraction to short-term memory (S/SD), but not further (L/LD), among the older subjects only.

(c) *Effects of the distraction task on the no-memory task among younger subjects*

The auditory distraction task had no effect on errors or speed in the no-memory task (N/ND) among the younger subjects. As will have been noted, an ND group was not formed when repeating the experiment with the older subjects.



*(d) Performance on the distraction task*

(i) While performing on the short-term memory task (SD) younger subjects averaged nine items correct on the auditory distraction task, but 6 of the 8 older subjects failed to record more than three correct items.

(ii) While performing the long-term memory task (LD), younger subjects averaged seven items correct on the distraction task, but 7 out of 8 older subjects failed to record more than three correct items.

(iii) In the SD group, 1 older subject did as well as most younger subjects in the same time, while another achieved the highest correct score on the distraction task by taking much more time.

(iv) In the LD group, 1 older subject did as well as most younger subjects by taking much more time.

*(e) Differences between younger and older subjects*

(i) All groups on tasks involving memory (S, SD, L, LD) show larger differences in speed between the old subjects than between the young ones (variance ratios,  $P < 0.05$ ). Typically the quickest subjects in both young and old groups are similar in performance; but the slower older performers are much slower than the slowest young ones.

(ii) Older subjects make fewer isolated errors of omission on the short-term memory task with distraction (SD). There are no significant differences between young and old in sequence errors (RCNS and RCF) on this task.

(iii) There are no significant differences between older and younger subjects on errors of any kind on the long-term memory task, with or without distraction.

(iv) While all younger subjects are able to achieve a fair performance on the auditory distraction task, most older subjects fail to do so: given two tasks to perform simultaneously they tend to drop one of them almost completely.

*(f) Individual differences*

In an effort to explain—or at least to hypothesize about—the fact that some individuals perform better than others in each group, the relationships between performance and available measures of intelligence and personality have been investigated. In the young groups, twenty-five such relationships could be evaluated: of these, four were found to be significant. In the old groups (for whom additional personality data were available) forty relationships with performance could be evaluated: of these, eight were found to be significant. Summarizing these, it can be said that differences in intelligence are associated principally with isolated errors of omission (high intelligence: few errors), and that personality differences are associated principally with speed of performance.

But these data have not helped us to account for individual differences in sequence errors under distraction (SD, LD) in the older groups. An alternative approach is to examine in detail those three older men in groups SD and LD whose performance on the auditory distraction task was comparable with that of the younger men.

(i) *Group SD*

(a) One man achieved the highest score on the auditory distraction task—considerably higher than even the best young subject. In doing so he made the second fewest number of sequence errors in the combined group ( $N = 14$ ), and his isolated errors of omission were average. But this was all achieved by taking the longest time (more than twice the shortest, and half again as long as the slowest young subject). This man is the oldest (55); he is of very high non-verbal intelligence (Progressive Matrices (PM) score 54 out of possible 60); has an average vocabulary (Mill Hill Vocabulary (MHV) score 25); is emotionally stable but socially introverted; and presents the least 'rigid' personality of the 8 older subjects. He left elementary school at 14 and has been an insurance clerk ever since.

(b) Another man obtained an auditory distraction task score equal to that of the second-best young subject. In his case this was achieved in average time, but at the expense of many errors, both isolated and sequence. This subject is above average age for his group (53); of high non-verbal intelligence (PM 50); possesses the highest vocabulary (MHV 31); is rather unstable emotionally, but socially extroverted; on the personality rigidity scores he was average for this group. He left technical school at 16 and has been a skilled pattern-maker ever since.

(ii) *Group LD*

Only one of the older men achieved a score on the auditory distraction task comparable with those of the younger men. He made fewer errors than was average for the combined group ( $N = 12$ ), but took the second longest time (nearly twice the shortest, and one-third again as long as the slowest young subject). This man is of average age in his group (52); of average non-verbal intelligence (PM 43); possesses a small vocabulary (MHV 19); is emotionally stable, but rather socially introverted; he obtained rather high scores for personality rigidity. He left elementary school at 14, and has been a policeman for the past 30 years.

## V. DISCUSSION

These simple tasks appear suitable for use in further research on problems of this kind: they were easy for the subjects to understand, and no especial difficulties were encountered during testing. Both memory tests gave rise to errors which can be firmly identified as due to failure of memory. (The 'short-term' memory test is easier to score and may be preferable for that reason.)

We have been able to show fairly clearly that search for a signal when the subject has to *remember which signal is required* is likely to be severely affected by a 'distraction' task which involved intake of subsidiary but still significant information. This finding seems to support the view that immediate memory involves continuous action on the part of some mechanism which is also needed for perception, and that this mechanism cannot handle both tasks simultaneously. It is, in other words, sensitive to what has been called 'interference'.

It has also been shown that this limitation is more marked in older than in younger people. Most of these subjects in their fifties appeared quite unable to perform two

tasks of this nature simultaneously without either making many errors on the main task, or taking much more time over it. The fact that a few older subjects can, by taking much more time over the main task, achieve a reasonable performance on the 'distraction' task *without* increasing their errors on the main task is, however, a little puzzling.

It could perhaps be argued that, if a man is prepared to rehearse repeatedly items which he wishes to remember, those items are transferred from short-term to longer-term storage and so become very resistant to interference by distraction. Such a growth of resistance through allowing opportunity for rehearsal has been shown by Sanders (1961). Thus those exceptional older subjects, who make few errors at either task, may be rehearsing to an abnormal extent each number and letter they have to remember. This time spent in rehearsal would slow them down at their cancellation task, but would prevent the forgetting of crucial information from either task while a signal from the other is being perceived.

It may seem curious that the time taken by, say, listening to one auditory letter should be enough to allow decay of the stored information from the main task. One of us (Heron, 1962) has, however, provided some evidence to suggest that among older subjects the 'interference' theory may be more relevant in considering performance at tasks involving immediate memory than is decay theory—though both may be necessary. Since the subjects in that experiment (on dialling) were the same older men as those who served in the present investigation, comparisons are made easier than would usually be the case. Briefly, it was found in this dialling experiment that with serial auditory presentation of eight-digit messages these older men did not show the expected difference between performance at immediate rehearsal and at immediate dialling. The suggested explanation for this finding was that 'the arrival of another digit in a series may constitute interference in older subjects so great and so cumulative that in a series of eight digits the effect is enough to make the dialling delay relatively unimportant'.

In the present experiment the arrival of random letters from the distraction task may have had its main effect upon older subjects through interference rather than through decay.

## VI. CONCLUSIONS

1. Continuous tasks which involve even a slight load on memory are very vulnerable to distraction, when compared with similar tasks which avoid this load.
2. In such tasks old people differ markedly amongst themselves, some being seriously inferior in performance to younger people, while no corresponding difference in variance appears in a similar task without memory load.
3. When a distracting task is presented together with a main task involving memory, old subjects do very badly at one or the other, whereas young subjects do reasonably at both.
4. Some features of the results emphasize the importance of interference rather than solely decay in short-term memory, and of rehearsal as a possible means of combating interference.



This paper was originally presented at the London Conference of the British Psychological Society in December 1960.

The authors wish gratefully to acknowledge the assistance, both with experimentation and with data analysis, of Mrs Margaret Gregory in Cambridge and of Mrs Brigid Brechling and Miss Margaret McEwan in Liverpool.

#### REFERENCES

- BROADBENT, D. E. (1957). Immediate memory and simultaneous stimuli. *Quart. J. Exp. Psychol.* **9**, 1-11.
- BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quart. J. Exp. Psychol.* **10**, 12-21.
- CONRAD, R. (1957). Decay theory of immediate memory. *Nature, Lond.*, **179**, 4564.
- CONRAD, R. (1958). Accuracy of recall using keyset and telephone dial, and the effect of a prefix digit. *J. Appl. Psychol.* **42**, 285-8.
- FRASER, D. C. (1958). Delay of immediate memory with age. *Nature, Lond.*, **182**, 1163.
- HERON, A. (1962). Immediate memory in dialling performance with and without simple rehearsal. *Quart. J. Exp. Psychol.* (to appear).
- INGLIS, J. (1960). Dichotic stimulation and memory disorder. *Nature, Lond.*, **186**, 181-2.
- KAY, H. (1953). Experimental studies of adult learning. Ph.D. Thesis, Cambridge University. Experiment described in Welford, A. T., *Ageing and Human Skill*. Oxford University Press, 1958.
- SANDERS, A. F. (1961). Rehearsal and recall in immediate memory. *Ergonomics*, **4**, 29-34.

(Manuscript received 11 August 1961)

## PUBLICATIONS RECENTLY RECEIVED

*The Chosen Few.* By W. D. FURNEAUX. Oxford: University Press, for the Nuffield Foundation. 1961. Pp. xxvi + 245. 25s.

The decade following the end of the war saw the development of a bitter social, political and educational controversy over the pattern of organization of secondary education. The focal point for the attack against the tri-partite system embodied in the Education Act of 1944 was the '11 plus'. Criticism came not only from the extreme left, who saw the examination as one which favoured middle-class children at the expense of those with poor cultural backgrounds, but also from the extreme right, who found the emphasis on innate brightness a threat to the slower but surer industrious children of ambitious parents. Sociologists criticized the examination as a hindrance to social mobility, while educationists deplored it for its severe backwash effects on the primary school. During the past five years the controversy has begun to die down, partly as a result of the development of secondary modern schools, and the demonstration by a substantial (and growing) number of them that G.C.E. success is a possible goal for more than an insignificant minority of their pupils. As the '11 plus' recedes from the centre of the educational stage, the '18 plus' seems likely to take its place in the limelight with essentially similar criticisms being mounted from the same sources. The unprecedented increase in the size of sixth forms, and in the entries for G.C.E. O-level and A-level subjects, is producing larger and larger numbers of pupils qualified for, and desirous of, a university education. It is clear that the planned expansion of British universities cannot hope to keep pace with this increase in applicants, so that the pressure on universities, and on candidates—already severe—will grow considerably in the next decade.

Inevitably the process of selection will increasingly become the focus of critical examination, and the universities will find it more and more necessary to assure the general public that their selection methods are equitable and efficient. Comparatively mild pressures of this kind have, of course, existed for a long time, and all universities are aware of them. But it is doubtful whether they all appreciate the potential build-up in public criticism, as more and more technically qualified applicants, from all kinds of schools—modern, comprehensive, grammar and public—find it impossible to secure admission. The need to provide a demonstrable efficiency in selection, as opposed to an assumed efficiency, is hardly recognized as yet.

Following the 1947 Conference of Home Universities the Nuffield Foundation allocated money for a 10 year investigation of the selection problem. The plan of research included the application of psychological tests, not only of ability and aptitude, but also of personality and temperament, so that we might discover whether the type of selection well-known in the U.S.A. might be usefully used in this country. Mr Furneaux's book represents the first volume of the final report on this major investigation. It is, however, severely limited in scope, since it does not include any data on the thirty or more psychological tests used in the research. Nevertheless, it is an important and valuable report, covering aspects of selection and determinants of success which are of considerable interest to teachers, university selectors and psychologists.

Mr Furneaux starts by considering the path to the university through the secondary school, from '11 plus' through G.C.E. O-level to G.C.E. A-level. He is particularly concerned about the reserve of talent, or pool of ability, and therefore discusses such questions as the potential in the secondary modern school, the premature-leavers in the grammar school and the standards of the G.C.E. examinations. He produces some most interesting tables, tracing the passage of 5000 boys and 5000 girls from primary schools through their secondary education, estimating what effect might be produced by eliminating premature leaving, and by increasing the size of grammar school intakes. He concludes that nearly twice as many boys, and nearly three times as many girls could eventually reach the stage of applying for admission to a university.

He has an interesting chapter on the effects of multiple applications. The average correlation of universities' judgements on candidates is about 0.4. He concludes—as we might expect—that multiple applications improve validity. Misclassifications by one university are reduced by nearly 40% if application is made to a second university. Adding a third application produces a further drop of 12%. The increase in validity is, however, only achieved at the cost of 'a very considerable expenditure of effort'. It is clear, too, that the more applications a candidate makes, the better his chances of getting in.

In considering socio-economic factors Mr Furneaux concludes that university selection really begins at birth, for 'a child's academic history is strongly influenced by the social class into which he is born'. He suggests that socio-economic factors cease to operate once the pupil enters the sixth form—they have no effect in determining whether he will be accepted by a university. It does not, of course, follow that if the selective effects of occupational-group were removed in the preceding stages of education, they would not then be reintroduced at the stage of university entry as at present organized.

It is clear that an evaluation of Mr Furneaux's results depends upon the representativeness of his sample. The testing programme (not yet reported) was carried out on students at the University of Sheffield (1948–55). Many of the data in the present volume were obtained from a sample of contributing schools. The 81 L.E.A.s responsible for all schools in the country from which more than five pupils had applied to Sheffield for admission in 1948 were asked to co-operate. Sixty-nine agreed, and from the 600 schools administered by them, 280 agreed to become fully co-operating. By 1953 some 90 of these were dropped out of the programme because of the difficulties caused by the size of the testing programme. So it will be seen that we are dealing with a sample firmly based on one provincial university, and that the school sample was far from complete, and far from random. This is not a criticism so much as an emphasis on the limitations of interpretation. These are signs that the school sample may have been a curious one, particularly for co-educational schools (see Table 5.4), and it must be recognized that comparisons such as Arts *v.* Science may be affected by the strength of, and the predilections of applicants for, particular departments in a particular university.

The major doubt one has about Mr Furneaux's conclusions stems from the difficulty inherent in a report of 10 years' work in a field where temporal trends are very strong. In considering the 'reserve of talent' one is worried to find the degree of dependence on the Ministry's report on 'Early Leaving', published in 1954. Since that time an unprecedented increase in sixth form numbers has occurred. In 1958 there were 76,000\* pupils in sixth forms in England; by 1960 this figure was 99,000.† The N.U.J.M.B.'s O-level entries in 1959 showed an increase of 50 % in Arts subjects, 80 % in Mathematics, 110 % in Chemistry and 120 % in Physics over 1951.‡ G.C.E. entries for the whole country from secondary modern schools rose from 10,000 in 1958 to 22,000 in 1960.§ One feels that Mr Furneaux has made some allowances for the 'bulge' (in birthrate) but little for the 'trend' (in sixth form proportion). The hazards of extrapolating from existing trends are obvious, but must be made by anyone attempting to estimate the reserve of talent.

It is clear that this is an important book on a highly significant topic, and it will be welcomed by educationists, sociologists and psychologists. Whether it will be equally welcomed by university administrators and those departmental heads responsible for selection procedures is more doubtful. The existing inefficiency of the average university's methods of keeping records is well brought out by Mr Furneaux. He might have argued even more strongly not only the importance but the duty which universities have of making their own evaluation of the efficiency of their selection methods. If this were to become an accepted practice, then record systems would rapidly be re-organized.

STEPHEN WISEMAN

*Perception and the Physical World.* By D. M. ARMSTRONG. London: Routledge and Kegan Paul. 1961. Pp. xii + 196. 25s.

*Sensation and Perception: A History of the Philosophy of Perception.* By D. W. HAMLYN. London: Routledge and Kegan Paul. 1961. Pp. xi + 210. 25s.

These volumes are among the first of the new International Library of Philosophy and Scientific Method, under the editorship of A. J. Ayer. They succeed an earlier series which, under C. K. Ogden, gave us many works that have since become classics. The dropping of the word 'psychology' from the title of the new series is, I presume, a recognition of the autonomy which our science has since achieved, but it is to be hoped that psychologists will not on that account ignore the series and so contribute further to a rift that can only impoverish both sides.

The authors are both young philosophers thoroughly at home in the post-Wittgensteinian

\* *Education in 1958.* H.M.S.O. 1959, table 11.

† *Education in 1960.* H.M.S.O. 1961, table 11.

‡ Wiseman, S. (Ed.), *Examinations and English Education 1961*, p. 46.

§ *Education in 1960.* H.M.S.O. 1961, table 43.



climate of present-day philosophy. Moreover they approach perception from an essentially similar standpoint, although in Hamlyn's book this is largely implicit in his criticisms of historical theories, whereas Armstrong is putting forward an explicit thesis. The reader would therefore be advised to start with the Armstrong.

The question with which Armstrong confronts us at the outset is: 'what is the direct or immediate object of awareness when we perceive?' He maintains, rightly I think, that philosophers can be divided into three schools according to the answer which they give to this question; the three schools being Representationalism, Phenomenalism and Direct Realism. One of the reasons that has made perception such a baffling problem for philosophy is that grave objections can be levelled against all three doctrines and yet between them they appear to exhaust the alternatives! Armstrong, however, elects to defend Direct Realism which he defines as the doctrine that 'the immediate object of awareness is never anything but a physical existent'. Direct Realism might also be described as the sophisticated successor of Naïve Realism and on p. 139 Armstrong claims that it is 'this simple and straightforward view of perception and the physical world that has come off unscathed' from his analysis. However in the final part of the book *The Argument from Science*, he proceeds to make so many concessions to potential critics that hardly anything is left of this claim.

Nevertheless, this is perhaps the most thorough attempt yet made to state the case for a theory of perception which for many contemporary philosophers has come to assume almost the status of an 'official line' and I must say that the case is put more strongly than I would previously have thought possible. Armstrong, moreover, always reveals a willingness to take philosophical ideas seriously, and this by itself makes his book a refreshing contrast to the weary sophistication and heavy pedantry of so much current philosophical writing.

Probably the most original feature of his treatise is his ingenious attempt to identify perception and belief. As opposed to the conventional Empiricist view that sense-experience affords the ultimate grounds on which our beliefs about the world are based Armstrong repeatedly insists that sense impressions are *nothing else* than our beliefs about particular physical objects, beliefs which are true in the case of veridical perception, false in the case of illusions; no question of any sensory content of the perception arises. I have no space to discuss the implications of this very bold proposal but I wonder if Armstrong fully realizes how paradoxical some of them are? The trouble is that whenever he finds himself in deep water he browbeats us into acquiescence by reminding us that the only alternatives are Phenomenalism or Representationalism, both doctrines which he claims to have conclusively refuted. But has he? As regards Phenomenalism, it is true that I found myself in almost complete agreement with his criticisms although even here he shows it to be not so much false as intolerably paradoxical. But as regards the Representative theory, he so much takes for granted that it has already been demolished that he accords it only the most cursory treatment. Yet the objections he brings against it carry weight only for those who are already committed to a dogmatic positivism. Not being one of them (and, to judge from his remarks on p. 159, neither is the author) I found myself continually answering his rhetorical questions in the opposite sense from that intended.

One problem that confronts a Direct Realist is the relationship between sensation and perception. For, if in perception we are in immediate contact with the external physical world, perception cannot also consist in having sensations. Where then do sensations come in? According to Hamlyn (p. 196), 'it shows a misunderstanding of the problem if one is tempted to ask what experiences, processes or activities constitute perception'. But he adds 'this is just the question which has so often been asked in the history of philosophy and psychology'. And successive philosophers are here taken to task for assimilating the concept of perception to that of sensation. Yet on the correct relationship between sensation and perception that, while it is all right to say vague. He is understandably dissatisfied with Rylo's suggestion that when we see things we have that when we feel things we have touch sensations, to suppose that when we see things we have visual sensations is to perpetrate a para-mechanical myth. He admits, moreover, that sensations are necessary to perception, but how to bridge the gap between the subjective sensations and the objective perceptions is something that never quite emerges from the tortuous reasoning of his final chapter.

As a history, however, this is an impressive performance, being at once erudite, readable and compact, even though it suffers from something of the same negative approach as the author's previous *Psychology of Perception*, where contemporary theories were the target of his criticism.

JOHN BELOFF

*The Growth of Basic Mathematical and Scientific Concepts in Children.* By K. LOVELL. London: University of London Press. 1961. Pp. 154. 15s.

This book gives a lucid exposition of current thinking on the formation of basic mathematical and scientific concepts. It brings together the results of the extensive researches of the Piaget school and of the many studies which have been carried out in order to verify or check their findings. Much work has been done in this field by Dr Lovell himself, either on his own account or in collaboration with his students, who have carried out some seven thousand experiments of the type first suggested by Piaget or Inhelder.

There are chapters on the possible stages through which children pass in acquiring concepts of number, substance, weight, time, space, length and measurement, area and volume. Two chapters are devoted to an exposition of the rationale of various methods of teaching number concepts such as those associated with the activity and environmental schools or with the structural approaches of Stern, Cuisenaire and Dienes. These chapters, and one in which the concept of number is discussed from the mathematician's point of view, should be helpful to all who, perhaps themselves not mathematicians, may be concerned with the teaching of mathematics in infant and junior schools.

While the studies reported and discussed by Dr Lovell agree in broad outline with the findings of Piaget and his collaborators, much evidence is advanced that sheds some doubts on the correctness of Piaget's theories in their entirety. As an example we may take the general thesis that the child's conception of space begins with topological concepts, which are transformed concurrently into concepts of projective and Euclidean space. There appears to be little evidence to suggest that it is topological relationships as such which enable a young child to identify some shapes more easily than others. It is more likely that it is 'gaps, holes, curves, points, corners, ins and outs etc. in Euclidean space that make identification easy, because the amount of "information" conveyed is greater'. Dr Lovell emphasizes, however, that these criticisms do not detract from the value of the ingenious experiments devised by Piaget and Inhelder.

The book is well documented, and each chapter includes a comprehensive list of up-to-date references.

I. MACFARLANE SMITH

*Psychology of Judgement and Choice: A Theoretical Essay.* By FRANK RESTLE. New York and London: Wiley. 1961. Pp. xiv + 235. 56s.

The task the author has set himself is to examine the logico-mathematical procedures commonly employed in many branches of psychology, chiefly those concerned with probabilistic phenomena. In the first part of the book he introduces the reader to set theory and in the second part he illustrates the use of set-theoretic concepts and methods in the study of utility, recognition, prediction, gambling, reaction time and the determination of thresholds.

His discussion of these various topics is integrated by his concept of choice, which he interprets as involving the comparison of any new situation with the 'schemata' of other situations in previous experience. This point of view raises the question of *access* to the 'schemata' in memory, a question which is profitably discussed in later chapters.

He is anxious to make his general position clear. It is neither that of the introspectionist, which he believes is committed to metaphysical dualism, nor that of the behaviourist, which he considers sterile. He takes what he calls a 'cognitive' view which interrelates physical and mental descriptions in a single conceptual framework. This avoids having to deal with them as the abstractions from reality which they would be if studied separately. In taking up a problem he begins, like the Gestalt and other theorists before him, with qualitative distinctions, which he then proceeds to describe in terms of sets and the *measure* of sets.

Having praised the book highly, it will not, I trust, be taken amiss if I make two critical comments. First, the author, as is not unusual among American psychologists, seems rather out of touch with non-American sources which are directly relevant to his problems, particularly (but not solely) Professor G. L. S. Shackle's important work on choice. Secondly, the author's flair for going into theoretical orbit often creates landing difficulties when he proposes to return to empirical *terra firma*. When all is said, however, the publication of his book will lift him above the rank and file of psychologists and go far to establish a well-merited reputation for high competence.

JOHN COHEN



*Markov Learning Models for Multiperson Interactions.* By P. SUPPES and R. C. ATKINSON. Stanford University Press; London: Oxford University Press. 1960. Pp. xii + 296. 66s.

This book attempts to combine three recently developed disciplines: the behaviour of small groups, the study of learning as a stochastic process, and the theory of games. It represents one of the many developments from Estes' well-known article in the *Psychological Review* of 1951, entitled 'Toward a Statistical Theory of Learning'. The authors ask themselves how far learning theory can be extended to predict behaviour in situations that correspond closely to real games, such as guessing which light is switched on next. The experiments described vary in the amount of information given to the subjects concerning other persons' responses, a restriction which turns the subjects into learners as well as players. It is maintained that when individuals reach an asymptotic level of behaviour (in learning-theory terms) they are in equilibrium with their environment, hence the optimality concepts of game theory also apply to their patterns of choice. These experiments can thus be regarded as a kind of competition between game theory and learning theory in their relative ability to predict behaviour.

The learning axioms adopted by Suppes and Atkinson assume that there is a fixed number of responses and reinforcements and a fixed set of stimulus elements for any specific experimental situation. The number of stimulus elements is assumed to be small. Hence the authors are able to consider explicitly the appropriate Markov chains and processes derivable from the theory. The basic modification compared with the Estes-Burke sampling theory is the conversion of the conditioning process itself into a probabilistic instead of a deterministic process, one of the conditioning axioms assumed being that if a stimulus element is sampled on a trial it becomes conditioned with probability  $\theta$  (not unity) to the response that is reinforced on that trial. It is claimed that such generalized conditioning models receive their first systematic treatment in this book, the use of Markov processes being justified on the ground that they provide a framework within which the sequential aspects of behaviour can be analysed. The authors contend that, from the systematic standpoint, a theory or model based only on qualitative distinctions leads to a small number of testable predictions, and quote Aristotle's physics and Lewin's topological field theory as examples of this. Earlier attempts at quantitative theory did not, in their view, lead to a theory that was mathematically viable, e.g. in Hull's theory it is impossible to make non-trivial derivations leading to new quantitative predictions of behaviour. In statistical learning theory we have, on the other hand, a theory with the same sort of 'feel' about it that theories in physics have. Non-trivial quantitative predictions can be made. Once we make an experimental identification of stimuli and reinforcing events it is usually clear how to derive predictions about responses in a manner that is not *ad hoc* and is mathematically exact.

The one-stage Markov model is clearly limited, since it assumes that the subject pays attention to the results of the previous trial alone. The authors find that models with one stimulus element and a conditioning parameter  $\theta$ , do not yield realistic predictions concerning the fine structure of the data, although they do a surprisingly good job of predicting mean response probabilities. Thus the one element model is best viewed as a simple device for computing the grosser predictions of the general theory. In complex situations, however, this model can be used only for exploratory work, because of the enormous number of algebraic and numerical computations involved, and substantial applications have been made only in respect of two-person games and the effects of monetary pay-off. In these experiments the authors may claim some success in reproducing the fine detail of the data, although many psychologists would regard the primary aims of theory to be rather economy of thought and experimentation. The generalized conditioning model is also valuable in clearly indicating the differential effects of particular reinforcement schedules and the amount of reinforcement. The results in respect of game theory are more decisive; there is little, if any, evidence that it is an effective descriptive theory of actual behaviour.

The book is admirably written at a high technical standard, but the results are rather depressing. Despite the highly controlled experimental situations, the rigid theoretical framework and the complex methods of sequential analysis, no one really knows what  $S$  is going to do next.

F. W. WARBURTON



*Tactics of Scientific Research.* By MURRAY SIDMAN. New York: Basic Books. 1960.  
Pp. x + 428. \$7.50.

This book on scientific research is divided into four sections, dealing with theory, replication, generality and experimental design. In the past, problems of reliability and validity have been regarded mainly as the concern of educational and vocational psychologists, and the present work breaks new ground in examining these topics from the point of view of experimental rather than applied psychology. The author worked originally with Skinner at Columbia, and most of the experiments he describes in illustration of his argument are studies of learning in the free-operant situation. The position he adopts goes far beyond his earlier protests against 'average' functions, for he now takes the view that the study of groups and the study of single persons belong to different sciences. He illustrates his thesis by examining the slopes, intercepts and linearity of the curves of performance of individual subjects, and reiterates that these are commonly ignored in the process of simple averaging. He does not imply, however, that the study of the individual will lead to precision outside the laboratory, and he points out that no one expects the physicist to study gravity by observing the fall of autumn leaves. But Sidman's major concern is with variability. Secure within his laboratory walls, he attempts to tame apparatus, which ultimately transform the sources of variation into independent variables. He does not favour what he would regard as the methods of appeasement commonly adopted in other branches of psychology, e.g. the use of large samples and the application of complicated statistical techniques, which do not actually remove variability in the data. To the author, chance is simply a name for the combined effects of uncontrolled variables. If these variables are known, then a reliance on random action and chance is simply an excuse for incompetence. If the uncontrolled variables are unknown, then chance is a synonym for ignorance. Psychometrists will envy Sidman his optimism about both variability and the study of the single individual, although this may well be justified in the field of animal learning. On the other hand, they have the consolation that the complexities of learning by human beings have led to the development of techniques far more sophisticated than any suggested by Sidman. Experimentalists are an unconscionable time appreciating the power of factorial and Fisherian techniques or the value of theoretical work on variability and unreliability. It is rather quaint, for example, to read the view that if an experiment is performed with a single organism as the subject, intersubject replication *may* be demanded on the ground that the original subject could have been a freak, or to be told that psychologists are not as yet in substantial agreement with respect to which of the multitude of variables that affect behaviour will prove to be most efficient as a framework for systematization, e.g. that variables such as 'reinforcement' and 'extinction' may be combined to yield the broader behavioural process (i.e. factor) that we are accustomed to call 'discrimination'.

On the whole, this book is written at a general rather than a technical level, although to the specialist the detailed description of techniques in the last section will be invaluable. The remainder of the book will make interesting light reading for the statistical psychologist.

F. W. WARBURTON

*The Methods and Findings of Experiments on the Visual Discrimination of Shape by Animals.* By N. S. SUTHERLAND. Experimental Psychology Society Monograph No. 1. Cambridge: Heffer. 1961. Pp. iv + 68. 20s.

It is a pleasure to welcome the first monograph to be issued by the Experimental Psychology Society: the contents solid, straightforward and factual, if perhaps a trifle academic, the appearance unassuming and unexceptionable.

Dr Sutherland reviews evidence obtained over the last 50 years in the field of visual discrimination in animals; the monograph is presented as part of a project on stimulus analysing mechanisms and the present volume is confined within very narrow limits. The evidence included here is that concerned directly with the question of dimensions along which animals classify shapes. The author considers the evidence unsatisfactory, insufficient to provide secure generalizations about form discrimination in animals: therefore the early part of the book deals with purely

methodological problems. In this part are included sections on scoring methods, experimental situations, choice of species, training techniques, transfer tests and so on: it is argued that if these methods were followed correctly we could know what are the dimensions analysed by a given species, and research in this field could move on to consider neural mechanisms involved in discrimination.

The author excludes the discussion of theoretical problems except in passing, and several classes of experiment are also excluded as logically separate from the central problem, such as, for example, experiments on the effects of early experience, and also problems arising from innate patterns of behaviour. This restriction of the subject matter—probably inevitable in a short work—gives the presentation a feeling of some monotony. Nevertheless, the author has performed an immense service to those working in this field and in the field of discrimination learning by disentangling evidence obtained from many different methods, different stimulus patterns and a wide variety of animal species, classifying it, and making it available in one volume. There are more than 200 references and a subject index.

MARGARET VINCE

*Porpoises and Sonar.* By WINTHROP N. KELLOGG. Chicago: University of Chicago Press. 1961. Pp. xiv + 177. \$4.50.

To British psychologists, Winthrop Kellogg is best known for his pioneer work in rearing together a baby and a chimpanzee. It seems he has an eye for unusual research. For some years now he has been working with small whales, and in particular the bottlenose dolphin, the main concern of this book. If the addicts are to be believed (and it is unfortunate that most of us will be unable to get first-hand information ourselves), the dolphin is superior to the chimpanzee in intelligence; superior to song-birds in their communication by sounds; and superior to the bat (and the U.S. Navy) in sonar, or echo location.

The book does two things. It gives a general account of anecdotes about and observations and experiments on porpoises, from Pliny onwards. The account does not include the recent much publicized work of J. C. Lilly, but nevertheless it is enchanting, and written in a clear layman's English. The book also summarizes the experiments which have led, with the fascination of criminal detection, to the discovery that the incredibly accurate navigation of porpoises depends on a method of echo location. Kellogg has done many of the experiments himself, and he sometimes takes the descriptions word for word from his journal accounts. This part of the book is more technical, but it is fascinating to follow the discovery of such a beautiful mechanism. The bottlenose dolphin emits clicking sounds from the blowhole. The clicks vary in frequency and intensity, and also each click has a range of auditory frequencies resembling white noise, plus high ultrasonics. Kellogg believes that an analysis of the echoes of the clicks, in terms of time-intervals, intensity, frequency content and binaural comparison, provides the dolphin with information not only about distance and direction of objects in the sea, but also about the size, shape and even texture of the objects. (It is easy to see why the U.S. Government has been interested to finance much of the work.) On the perceptual side, it appears that the auditory mechanisms have evolved in a way which makes it plausible that the necessary analysis of echoes can be carried out.

The work reported in this delightful book adds one more to that already large group of animals (including man) that find useful information in the time-interval between self-made sounds and their echoes. That such time-intervals can be assessed with great accuracy may turn out to be a revealing aspect of nervous system functioning.

B. M. FOSS

*Die 'Sprache' der Tiere.* By FRIEDRICH KAINZ. Stuttgart: Ferdinand Enke Verlag. 1961. Pp. viii + 322. DM 26 (paper), DM 29.50 (bound).

The title of this book is slightly misleading. It is not so much the language of animals that concerns the author as the possibility that language proper (involving the construction of sentences) is the prerogative of man alone. As a result, he not only has to glide quickly and lightly over a large number of communicating systems in the animal kingdom but also has to deal with possible arguments from philosophy, archaeology, genetics, psychology, zoology, human physiology and anatomy, and sociology. The 300 pages of the text are therefore extremely tightly

packed, and to the English reader this in itself is somewhat offputting. Not only are we denied any form of obvious visual aid (a diagram every now and then would be most refreshing), but even the sub-headings are lost in the general text. There is no carefree skipping from point 1 to point 2 if the reader sees that point 1 is unlikely to interest him. Point 2 is tucked away in the middle of the paragraph and has to be winkled out. Even the appendices are mixed up with the end of the last chapter, and the bibliography is intertwined with a mass of notes on the text.

If the tendency of the publisher is to condense, that of the author is the opposite, and here again the reader accustomed to a different manner of presentation may become somewhat confused. Although Part 1 nominally deals with Biological Facts (and covers methods of communication from reptiles to apes) the author at any moment may indulge in a philosophical discussion on the nature of symbolism. Instances and examples are drawn from widely divergent sources and on a single page one may come across quotations from Aristotle, Bergson and Lorenz.

However the value of a work should not be judged on its lay-out, but on its contribution. The author has two main points to make: (1) animal communication systems are a 'vor-Gestalt' of human language and have developed the emotional gestures used by man in many situations; (2) the communication systems of animals are more stereotyped and less complex than those of man not because animals *could* not develop them more fully but because they do not *need* to. Just as the child whose screams are always satisfied by an indulgent parent may only learn to talk late in life, so the animal whose instinctive cries produce all it wants never has to develop more complex methods of seeking satisfaction. This is an interesting speculation, and it is disappointing that the section in which animal language is compared to that of human infants is restricted to two pages.

Interesting as these suggestions are, one cannot help wondering whether they are not based on rather flimsy evidence. A beginning has only just been made in the study of communication systems in the animal world, and it seems extremely likely that further detailed observations may show them to be far more complicated than hitherto suspected. It is not so long since von Frisch demonstrated what could be learned from the careful, systematic study of a single species and how dangerous it is to leap to conclusions based on preconceptions. What we want are more

MOYRA WILLIAMS

*Traumatic Aphasia. A Study of Aphasia in War Wounds of the Brain.* By W. RITCHIE RUSSELL and M. L. E. ESPIR. London: Oxford University Press. 1961. Pp. viii + 177. 38s.

Brain wounds in the Second World War were mostly caused by small fragments of metal from an exploding missile, which produced discrete, often cone-shaped, lesions; infection was rare because of early treatment. They could be localized relatively accurately, and provided good opportunities for studying the effects of brain damage on the mechanisms of language. This book reviews in a workmanlike manner the observations made in the 280 cases of aphasia in a series of 1166 cases treated in the Military Hospital for Head Injuries at Oxford. Special interest is taken in the cases in which the mechanisms of memory were affected.

The wounds were on the left side in 261 of the 280 cases of aphasia, in the mid-line in five cases, and on the right side in only fourteen cases. The left hemisphere is dominant for speech, therefore, as a general rule. After leaving out the cases in which there was evidence of damage to both hemispheres, there were only seven cases of aphasia in which the lesion was on the right side; four of these patients were left-handed, and three right-handed. Small wounds in the centre of the speech territory in the left hemisphere cause *central aphasia*, in which there is disorder of all aspects of speech; there is also mental confusion, and the post-traumatic amnesia tends to be long, perhaps because of injury to the fornix-hippocampal system. Special disorders of one or other speech function, such as motor aphasia, agraphia, or alexia, tend to result from small wounds on the periphery of the territory. Wounds in the posterior parietal area, on the periphery of the territory, which produce alexia tend also to affect a wide range of visual functions and thus produce severe loss of intellectual powers.

D. RUSSEL DAVIS



*Verbal Learning and Verbal Behavior.* Edited by C. N. COFER. New York and London: McGraw-Hill. 1961. Pp. viii + 241. 46s.

Here are the proceedings of a conference, the brain child of C. N. Cofer and D. D. Smith, which was sponsored by the Office of Naval Research and New York University, and took place in 1959.

'Verbal learning' refers to work in the Ebbinghaus tradition, and 'verbal behavior' to work which is in sympathy with, and an extension of Bartlett's view that neither should one believe that the nonsense syllable can divest the subject of his individuality, nor is it desirable to make the attempt. Under this heading the acquisition of language, and its influence on a person's interaction with the world around him are investigated.

At each session of the conference a previously circulated paper was discussed. Each of the seven main chapters of the book covers one session and consists of three sections: the original paper, a commentary on the paper prepared by the chairman of that session, and a brief summary of the discussion. Regrettably the procedure and presentation successfully insulate the reader from the liveliness of the debate. Moreover, the list of contributors evokes no surprise. This is a pity. The jaundiced view (admittedly it rests on homework done in Aberdeen on New Year's Day) is that the conference was too conservative in conception, and too anaemic at delivery.

But if the papers, which are summarized in the first chapter, are considered individually, one finds much that is of value. The three scholarly papers by Noble, Postman and Underwood, which are in the Ebbinghaus tradition, are well worth having. I like Russell's paper which raises a number of questions about the experimental investigation of verbal behaviour. But for me chapter 4 has most vitality. Bousfield, whom Osgood places 'on the side of Skinner and the angels', argues that meaning is an unnecessary concept in learning theory. Osgood mounts a lively attack, and wields his semantic differential vigorously, and, if you think it a well-forged weapon, effectively. I found much to disagree with in both papers, and enjoyed them. Goss, on the acquisition of conceptual schemes, disappoints. With his difficult style he is quite a problem.

As a conference on verbal learning and verbal behaviour this was not first class, but some of the contributions are. The reader will find much to applaud, and if he has borrowed the book from a friend he can have no grumbles.

J. R. SYMONS

*Fundamentals of Psychology.* By C. J. ADCOCK. Wellington: Price, Milburn and Co. 1959; London: Methuen. 1960. Pp. 220. 12s. 6d.

This is an eminently sane and scholarly introduction to psychology. One of its best features is its eclecticism which successfully integrates recent work with somewhat unfashionable, though important, earlier contributions such as those of McDougall and Shand. Dr Adcock, who is Senior Lecturer at the Victoria University of Wellington, New Zealand, writes as an academic psychologist who has sympathy with and understanding of the contribution of psychoanalysis to modern psychology. The book is less successful in its stated aim of presenting a development of the subject 'in which later chapters are very much dependent on the earlier ones' (p. 9), and the reviewer remains not wholly convinced by the claims of the dust-cover that 'any attempt to read chapters out of order will lead to difficulties'.

The text-book contains ten well-chosen illustrations, and its sixteen chapters are subdivided into seven sections. Following the introductory part, Section 2 deals with 'the primary sources of behaviour' (reflex systems, needs, and drives). Section 3 discusses the modification of these, through learning and conditioning. In Section 4 ('the affective reference frame'), the author includes an account of sentiments, and suggests that modern psychologists have been remiss in their neglect of this field. Section 5 ('the cognitive reference system') includes perception, perceptual learning, thinking and mental imagery. In Section 6 on 'individual differences' there is considerable emphasis on factorial work, notably that of Cattell. The concluding part (Section 7) deals with 'the patterning of personality', and places emphasis on psychoanalytic thought in general, and mental mechanisms.

Recommended references for further reading appear at the end of each chapter. In these, rather surprisingly for a book which exhibits obvious sympathy for both Freudian and Gestalt ideas, secondary references only are given. The text itself merits criticism in its neglect of certain information. Thus one can read about 'operant conditioning' (pp. 67-9; and again pp. 82-4),

without encountering the name of Skinner, an account of learning without reference to Hull, and of 'Gestalt psychology' without mention of Wertheimer, Kohler, or Koffka. Nevertheless the author has provided a useful introduction to psychology, balanced in its eclecticism, and admirable in its clarity and readability.

PETER MCKELLAR

*Human Behaviour. A New Approach.* By CLAIRE RUSSELL and W. M. S. RUSSELL.  
London: Deutsch. 1961. Pp. xii + 532. 42s.

It is hard to say to whom this profoundly odd brew will appeal; not, I suspect, to most psychologists, nor, one hopes, to many of the social workers, personnel managers, educationists, criminologists and others to whom it is also addressed. Dr W. M. S. Russell was a scholar in classics and English literature, who, after the war, turned to zoology and took up ethology under Dr Tinbergen (to whom the book is dedicated); his wife is a psychoanalyst of sorts—her own sort! All this, and a bit more, gets pulled in—the pivoting dance of the male zebra finch, the marital career of Henry VIII, neurology and the work of Halstead, the fixation of Faust, Sophocles and the King of Thebes, the cycles of Cathay, cybernetics, the European adventure, the mating behaviour of the male guppy, and plentiful snippets of case material. It is indeed, as the Russells admit, a 'great maze of material'; a bit like a bad film in which the director, lacking any sense of artistic proportion, throws everything in. It is a pity, because life is rich, and psychologists in their sympathies are often too narrow. The Russells with their diverse cultivation and literary gifts might have done a good job in getting psychologists out of their ruts. But there are so many silly, nebulous and uncritical statements in the book, and the basic ideology—a sort of Rousseauism—is so naïve, that such merits as it has are likely to pass unnoticed. Can we really believe that when Henry VIII's children were named their destinies were charted; that there is a correlation of 0.981 between authoritarianism and a preference for dogs rather than cats; that disputes about relative intelligence usually disguise disputes about penis size; that sexuality and intelligence are highly correlated; that the fully intelligent individual will be totally free of automatisms (have the Russells never read William James on habit?); that no clash between the interests of intelligent individuals is possible; that psychopaths are always the children of intensely respectable and law-abiding citizens; and so on, and so forth? We should certainly require a good deal more evidence for statements of this kind than the Russells provide in their book, and for their key concepts, such as 'pseudosex' (different altogether from 'real' sex), and 'masturbation fantasy' (of which the individual may have three, governing 'every aspect of life') much more precise definition. It is, of course, very appealing to dream euphorically of utopias in which 'the whole concept of power group vanishes from human history, and science and art become coextensive with human life... a world of eternally expansive happiness', and to believe that we can have it if only we intrepidly explore 'our own and each others automatic compulsions'. But it is possible that if we are to do this to much effect, and certainly if we are to build up something like a science of human behaviour, distinctly more discipline of the imagination—the sort of discipline which has proved necessary in all the other basic sciences—may be needed. And of this at present the Russells seem to have hardly an inkling.

L. S. HEARNSEAW

*The Moulding of Modern Man.* By T. H. PEAR. London: Allen and Unwin. 1961.  
Pp. 220. 21s.

This book is perhaps best described as a commentary on the contemporary scene by an acute and sensitive observer. Being a psychologist, Professor Pear brings the findings of recent investigations to bear on the problems, without neglecting other sources such as popular literature and the press. The discussion centres on 'modern encroachments on personal freedom', a topic essentially bound up with individual and social values; hence the author wisely refrained from adopting a pose of academic detachment, letting the reader know where he stands.

The first few chapters on self, personality and society prepare the ground; they contain some provocative speculations on the nature of self and its boundaries. One also encounters some incidental knocks at psychologists who confine their personality investigations to ticks and crosses on paper, obtained by remote control, or at small group experimenters who fail to consider the individual experiences of group members. The second part introduces reflexions on the

role of intellectuals, and the meaning and present implications of such concepts as loyalty and treason. This is followed by accounts of techniques of influence, particularly 'brainwashing', advertising and publicity; the effectiveness of these approaches, and popular reactions to them are considered. There is also a somewhat slight discussion of concepts of progress, and it was surprising to find, in a book dedicated to Professor Ginsberg, that the latter's brilliant analysis of this theme was not even mentioned.

Perhaps Professor Pear lays too much emphasis on the modernity of the modes of influence and persuasion he describes so well. The inquisition was not unfamiliar with techniques of thought reform, but recent studies of perceptual isolation help us to understand the causes of their effectiveness. Nevertheless it is certainly true that modern psychological knowledge, like all scientific advances, has potentialities for both good and evil, and it is valuable to have these implications so clearly set out. The book is written in an exceptionally lively, if somewhat discursive style, and is evidently addressed to the general reader. This need not deter the social psychologist, who can find many stimulating ideas in these pages, while aspiring Ph.D.s who pay careful attention to the numerous question-marks will discover a treasure house of possible topics for theses.

G. JAHODA

*The Self and Others.* By R. D. LAING. London: Tavistock. 1961. Pp. xii + 186. 25s.

This is Dr Laing's second book in which he demonstrates the manner in which the concepts of existential philosophy can be utilized to provide a fuller understanding of normal and abnormal mental activity. As in his first book the existentialist interpretation is combined with a phenomenological analysis of the observed data. Such an approach confines the author to observable and 'near' observable data and thus precludes inferential explanations which are an intrinsic part of psycho-analytic studies. As Dr Laing is seriously preoccupied with the relationships which exist between psycho-analysis, existentialism and phenomenology he presents a long and well reasoned statement regarding the phenomenological status of unconscious phantasy. This comprises the content of the first chapter of Part 1 of the book which is entitled 'Modes of Interpersonal Experience'. The second part of the book is entitled 'Forms of Interpersonal Action'.

This book has a number of striking characteristics. Above all it is remarkable for the number of philosophical and literary references. This is unusual in works which are primarily clinical in nature. Several of the literary fragments are presented with the purpose of illustrating the operation of the processes which the author suggests operate in interpersonal relationships whether healthy or diseased. The interpretations which Dr Laing makes of Dostoevsky's *Crime and Punishment* and Janet's *The Balcony* will, as the publisher's note on the dust-cover proclaims, excite the interest of most readers. The philosophical references have another purpose. The author borrows concepts from philosophers like Buber and demonstrates how they can be employed to obtain a more comprehensive view of interpersonal 'relatedness'. He shows the common ground which exists between Buber's view of the processes which operate between individuals and theories of certain American psychiatrists and sociologists who see subtle and complex disturbances of interpersonal relationships as being of decisive importance in the genesis of schizophrenic states.

This is not an easy book for those, like the reviewer, who are not well versed in philosophical literature and ways of thinking. Nevertheless, it is a book of interest because it represents the appearance of a new trend in the thinking of certain psychiatrists. Whether this approach will ultimately have any heuristic value remains to be seen. However, there can be no doubt that it will have an appeal which will extend beyond medical and psychiatric circles.

THOMAS FREEMAN

*The Living Symbol.* By GERHARD ADLER. London: Routledge and Kegan Paul. 1961. Pp. 463. 50s.

*The Springs of Creativity.* By H. WESTMAN. London: Routledge and Kegan Paul. 1961. Pp. 269. 45s.

From the titles of these books and the excellence of their production most readers of this Journal, probably, would surmise that they were contributions from the school of analytical psychology; examination of the pages, revealing such terms as animus, shadow, Great Mother,



Great Father, mandala, individuation, quickly confirms that first impression. Dr Adler's book, *The Living Symbol*, presents in detail a case of claustrophobia in a middle-aged unmarried woman which he considers to illustrate with exceptional clarity the stages of the individuation process. Mr Westman's, *The Springs of Creativity*, consists in effect of three essays: an introduction to Jungian concepts, interpretation of a number of biblical stories in accordance with these principles, and a case study of a psychotic or borderline psychotic girl whose treatment by Jungian methods he reports as successful. The problem of Dr Adler's patient is rooted in her disturbed relationship to a 'frightening' and emotional mother, leading her to take refuge in the intellectual sphere, with which her father was identified, and resulting in an impairment of spontaneous feeling, or, as Adler puts it, 'an alienation from the whole sphere of body and "earth", of eros and instinct' (p. 113). Mr Westman's patient, in the grip of a threatened schizophrenic breakdown, is less concerned with masculinity and femininity than with another of Jung's pairs of opposites, progression and regression, i.e. 'the force that strives toward wholeness, and thus toward creativity, and the inertia that strives to maintain chaos and thwart growth and creativity' (p. 118). The unsolved problems in both cases activate a wealth of symbols—archetypal images found similarly in mythological material; since the principle of compensation of homeostasis holds good at all levels of the organism, the psychic equally with the physiological, these symbolic formations point toward resolution of the difficulties, and when examined and discussed in the analytical situation bring about a wider integration of the personality.

Analytical psychologists will no doubt wish to comment on various points which amplify their own tradition, for example, Dr Adler's explanation that attempts to create the self are not only approximations, but approximations which are scattered and rebuilt time and again in the conflicts of life, a formulation which may bring this concept into somewhat closer relation with everyday occurrences. They will also refer to Mr Westman's biblical interpretations, a topic which the present reviewer is unqualified to examine, other than by raising a query, on historical grounds, concerning the feasibility of interpreting material from a relatively primitive society in accordance with more sophisticated concepts developed at a later time. For non-Jungians both works are specialist reading, to be tried, along with *Mythology of the Soul*, by H. G. Baynes, after the introductions by Frieda Fordham and Jolande Jacobi and some of Jung's own writings, *Symbols of Transformation*, for instance, and *The Practice of Psychotherapy*. For this group of readers the chief interest of both works can be expected to lie in the problems which they raise in common with all Jungian literature. What, for example, is meant by the archetype? Can 'the unconscious' use symbols in the way that analytical psychologists believe, as the best means of expressing ideas, often of a general nature, which up to that time are grasped imperfectly by consciousness? Or is the Jungian school dealing with pre-conscious material only, material which has previously been conscious and may become so again, as the Freudians have insisted? What are we to understand by individuation? Covering all these is a wider question: What has analytical psychology to bring to the body of generalizations and hypotheses concerning behaviour and experience which constitutes psychology as taught in the majority of universities and clinics, and how can the gap between the two be bridged?

Such questions can be touched on only briefly. The archetype, released from the mystification introduced by statements about the inheritance of ancestral experience, to which analytical psychologists are less prone than formerly, would appear to be an acceptable, indeed, a helpful, concept. That is, owing to the basic similarities of the human psyche and the conditions to which it is exposed at all times and places, it is not unreasonable to perceive broad similarities in the imagery of dreams and myths, as the Jungians continue to point out. From what mental level the archetypal material takes its origin is a more controversial question.

To the present reviewer it seems possible to bring analytical psychology into closer relation with the main body of psychological theory, e.g. by studying the part played by homeostasis and image-finding in emotional life and even subjecting them to some form of experimental scrutiny. Admittedly there remain the fundamental questions of origins and destiny before which human understanding ceases, and there remain certain value judgements—even, it may be, the judgement that an awareness of the ultimately mysterious is a necessary component of maturity. But at least a field for investigation can be cleared by making, rationally, demarcations such as these. We are indebted to *The Living Symbol* and *The Springs of Creativity*, as to all such studies, for reminding us that the intensive investigation of the individual case, as well as the examination of small samples of behaviour from many different cases, is a worthwhile form of research, and,

AVIS M. DRY

Ill.: Thomas: Oxford: Blackwell. 1961. Pp. xx+253. 68s.

D. H. STOTT

1960. Pp. xii + 479. 63s.

Psychoanalysis as the 'scientia' or 'proto-theory' behind psychotherapy has in the past been treated by Professor Eysenck as the Trojan Horse of Unscience smuggling the Deadly Armies of Superstition into the Celestial Scientific City. A scientist must be seriously faulted if he melodramatizes the conflict between opposing investigators and theorists. History should guard us against the methodological dogmatism of great systematizers, the 'disease of orthodoxy' of which the Webbs accused the Russian Marxists. The legacy of 'righteous indignation' and 'odium theologium' of pre-scientific controversy can haunt us in the computing room and laboratory as well as in the consulting room. The melodramatic mood is indeed endemic in an age when University dons still feel the need for crime fiction and television Westerns. Against this emotional time-lag of the culture, psychological analysis of the problems presented by abnormal behaviour makes little headway, for the solutions will depend not only upon the application of scientific knowledge, but upon what is equally necessary, at least, the cultivation of the scientific detachment, which has some kinship with the 'tragic sentiment'.

Is Professor Eysenck playing his lead-part less melodramatically these days? I think he is.

Many British psychologists would very willingly subscribe to the need for a rigorous general research programme of the neo-behaviourist type pursued in the mental health field, though not necessarily of the Hull-Pavlov variety he advocates here. The assumption that any all-embracing general theory of action or behaviour is in fact attainable is for most of us, however, no article of faith, but a provisional working hypothesis. We are pleased, in our traditional empirical pragmatism (and looking at the fate of the grandiose general theories so far proposed) to feel an absence of emotional 'engagement' to any of the edifices. We are, however, prepared to defend, with many 'tough-minded' young experimentalists,\* a more enthusiastic attitude to some of the 'miniature theories' devised in laboratory and computing room. We may even nourish a guarded hope that they may be extended: some of these studies encourage this hope.

Would not Professor Eysenck in this volume seem to be coming nearer to this cautious framework of expectations? I think so.

It would be a formidable task, to which few would feel equal, to review the vast general group-research programme on abnormal behaviour initiated by Eysenck in the *Dimensions of Personality* of 1947 and to give the present collection of important papers its proper place in the development. The plan can nevertheless be praised, even if the time has not arrived for a final assessment, for the way in which it has subjected, as these reports show, many Hullian-Pavlovian concepts to a wide variety of experimental tests. Certain it is that these papers are more serious and considered than any previous pronouncements from his theoretical standpoint had led one to expect. What we may call the 'Maudsley mood' has in the past contributed too much to the polarizing of psychological controversy and the disruption of scientific communication. Even the recent *Handbook of Abnormal Psychology* did not provide, as a 'handbook' surely should, for the Eysenck himself admits, not finally resolved. Nevertheless, we should exact no such self-advocacy of the learning theory approach. The style still seems a little too aggressively polemical to proselytise among the more empirically minded marginal voters. But we should, of course, be sorry (let us confess it) not to find him as readable at three guineas as at half-a crown. We should miss, and our students would miss still more, his stimulating if sometimes irritating shafts of 'tendenz-witz'. And when exposition is the task, his lucidity, his armoury of knowledge and reference (a burden he nevertheless carries with no loss of lightness of step and agility) are altogether enviable.

But for all that, there is no doubt in this reviewer's mind that Professor Eysenck is mellowing! What could be more removed from the 'odium theologicum' of chapter 12 of *Uses and Abuses* than (p. 14):

*'I would not like to be understood as saying that behaviour therapy has been proved superior to psychotherapy; nothing could be further from my intention. What I am claiming is simply that as far as they go—which is not very far—available data do not support in any sense the Freudian belief that behaviour therapy is doomed to failure, and that only psycho-analysis or some kindred type of treatment is adequate to relieve neurotic disorders. This Freudian belief is precisely this—a belief; it has no empirical or rational foundation. This Freudian belief is counter-belief, equally unsupported, to the effect that psychotherapy is doomed to failure and that only behaviour therapy is adequate to relieve neurotic disorders. What I would like to suggest is simply that a good case can be made out, both on the theoretical and the empirical level, for the proposition that behaviour therapy is an effective, relatively quick, and probably lasting method of cure of some neurotic disorders.'* (Italics, needless to say, by the reviewer.)

Apart from the statement that the Freudian belief 'has no empirical or rational foundation' (Eysenck seems content to equate this with *experimental* foundation) the statement seems quite unexceptionable, equally acceptable is the claim that behaviour therapy has an advantage in beginning, unlike psychoanalysis, with a theory of behaviour in *normal* animals and people, though one would dispute his assertion that learning theorists of different schools are 90 % in agreement and 10 % in dispute. Moreover, Eysenck goes on to say that he 'would be the first to acknowledge the tremendous service that Freud has done in elucidating for the first time some of these dynamic relationships (sc. "in respect of neurotic and psychopathic behaviour") and in particular in stressing the motivating role of anxiety'.

\* E.g. D. E. Broadbent: *Perception and Communication*. London: Pergamon, 1958, chap. 12.



Why in face of such urbane and generous admissions must it be concluded that the major theoretical papers in this collection (Eysenck's *Learning Theory and Behaviour Therapy*, and his *Modern Learning Theory*, Shoben's *Psychotherapy as a Problem in Learning Theory*, Wolpe's *Reciprocal Inhibition Theory*) all fail to establish a sufficiently cautious methodological analysis of the problem? The urgent need to develop techniques of changing abnormal behaviour based upon proper evidence and success and failure is not denied by his theoretical opponents. The Freudian model cannot, however, be faulted out of hand on the charge that it lacks 'an intelligible, objectively testable "modus operandi" which can be experimentally studied in the laboratory'. It seems more than likely that the difference between normal and abnormal human behaviour only appears critically in respect of a limited number of extra-laboratory social relationships. Moreover, the relation of theory to practice is a complex one. The implicit assumption of so much of Eysenck's attack on 'unscientific' theorizing is that before experiments are done there is no reliable knowledge. Yet we all know that many techniques were long ago developed of bringing about the unlearning of 'normal', and the learning of 'abnormal' behaviour in the domestication and training of animals. Eysenck underestimates the importance of this fact. These techniques are communicated, surprisingly, by the use of the concepts of a common-sense psychology of consciousness anthropomorphically attributing to animals 'will', 'desire', 'love', 'hate', 'aggression', 'dispositions' and the like as well as the normal 'subjective' calculus of 'pleasure' and 'pain'. Some fundamental relationships are obviously communicable in this way. It is our normal technique for the 'education' of humans too. The suspicion of most psychologists, moreover, is that the subtleties of this so-called 'mentalistic' language provide a fantastically complex stimulus-system, the use of which in the classical psychotherapeutic situation is an exceedingly difficult skill. That only a few really skilled practitioners are any more successful than a kind friend may very well be true. That no satisfactory theory is yet available to direct the creative therapeutic art should surprise us no more than that there is no behaviour theory which will enable us to teach a man to make jokes. Perhaps Ellis's\* suggestion of a 'rational psychotherapy' might lead us to reconsider whether we really can pension off 'the language of consciousness' now that we have learnt, ingeniously, to substitute in a number of cases a bit of desirable conditioning for an undesirable symptom.

Perhaps more to blame after all than Professor Eysenck (in this present temper), for the lack of proper discussion of the theoretical problem, is the sectarian dogmatism of those psychotherapists who would deny all long-term effectiveness to behaviour therapy even in cases where, for example, enuresis is a single main symptom, or in simple phobias. There is ample evidence in the examples given here and in the plentiful literature cited that behaviour therapy is, indeed, likely to take an important place with the physical therapies, chemotherapies, surgery, sociotherapy, psychotherapy (in its many different forms, for it is convenient, but naïve, to attempt to establish a single 'basic' anti-scientific devil in Freud) and the common-sense or 'rational' psychotherapies of Ellis and the kind aunts and wise uncles, and all. It is not satisfactory, either, to turn the tables by Eysenckian arguments against the admittedly uncontrolled and case-history illustrations which are advanced in many of the later papers. Eysenck himself sees that they are primarily illustrative despite the quantitative precision of, for example, the Yates paper on the extinction of tics in a female patient. There is now a vast amount of convergent evidence of the sort here advanced to support the general thesis that the psychotherapeutic situation is in essence a learning situation in which primarily verbal stimuli, and emotional and cognitive responses, are involved. Without a doubt certain cases with motor symptomatology may be advantageously, and possibly more precisely, analysed in terms of behaviour theory. Yet at the present stage of progress something simpler than Hull's over-ambitious hypothetico-deductive theory or Gwynne Jones and Eysenck's variant of it is required. Something more like Skinner's positivism seems likely, to this reviewer, to be more useful, but all power to *all* the theory-builders and theory-testers of whatever persuasion! History will sort them out, for theorists have to earn their keep!

G. A. WESTBY

\* A. Ellis (Ed.). *What is Psychotherapy?* New York: American Academy of Psychotherapists, 1959.

*Psychotherapy of the Psychoses.* Edited by A. BURTON. New York: Basic Books. 1961. Pp. x+386. \$7.50.

This volume of essays on the psychotherapy of the psychoses is a book for the specialist within a speciality. It will be of greatest interest to psychoanalysts and psychiatrists who are working in mental hospitals and adopting a psychotherapeutic approach to their patients. It is unfortunate that the editor has limited the contribution on manic-depressive reactions to a single chapter while thirteen are devoted to schizophrenia. This tends to bias the book towards the category of those concerned with the psychotherapy of schizophrenia. The majority of the contributions come from the United States, only three emanating from Europe.

Dr Burton has succeeded in bringing together the main current trends in the psychotherapy of schizophrenic states. While it is true that the classical psychoanalytic approach to these conditions does not appear in this book, there are a number of essays which do not deviate far from this model. These essays, contributed by workers from the Chestnut Lodge Sanitarium (Burnham, Searles, Stierlin and Will), represent the psychoanalytic position and they are characterized by a freedom from dogmatism and complete integrity. Mental hospital psychiatrists will recognize in the description of the clinical material their own every-day experiences. Another psychoanalytic contribution comes from Pious, who asserts that a new set of concepts is necessary for both the theory and technique of the psychotherapy of schizophrenia. Serious students of the psychoanalytic theory of schizophrenia may reflect that Pious has merely provided a new set of concepts to describe phenomena adequately covered by the hypothetical model advanced by Freud.

The Jungian approach to the psychotherapy of schizophrenia is well represented in a contribution by Perry, and a combined Jungian-Existential view is provided by Burton. Two quite new techniques are described by Don Jackson and Gisela Pankow. The reader is left with the deep impression that psychotherapy has an important place in the treatment of schizophrenic states. He will find it difficult to deny their efficacy even in the chronic case. Nevertheless he may justifiably consider whether the methods described in this book can ever be applied on a wide scale. Each method is time-consuming in the extreme and demands a high degree of training for the therapist. However, no one will doubt that through these techniques important new data about psychotic states will come to light.

THOMAS FREEMAN

*Prevention of Mental Disorders in Children.* Edited by G. CAPLAN. New York: Basic Books; London: Tavistock. 1961. Pp. xii+425. 42s.

*An Approach to Community Mental Health.* By GERALD CAPLAN. London: Tavistock. 1961. Pp. x+262. 25s.

Preventive psychiatry and the promotion of mental health among people in a community are the common, and indeed central concepts, of both books. The first, subtitled 'Initial Explorations', is based upon the discussions that took place at a conference held in Massachusetts in February, 1960. This conference was organized by the International Association for Child Psychiatry and Allied Professions as part of its preparations for the Fourth International Congress of Child Psychiatry to be held in 1962, the theme of which will be 'The Prevention of Mental Disorders in Children'. Sixteen speakers were invited to prepare preliminary draft papers, and in the light of the discussions at the preparatory conference the original papers were re-written. They form the main body of the book together with a general introductory and a concluding chapter by Dr Caplan. Only 'primary prevention' is considered, which is defined as 'the promotion of mental health and the lowering of the risk of mental disorder in a population of children by interfering with pathogenic forces within the children and in their biologic, psychologic and social environment before the appearance of identifiable pathology'. It is suggested that focusing upon the community rather than upon individual children introduces a new viewpoint. The first four papers deal with organic and pregnancy complications as etiological factors in emotional disorder; then follow four chapters on various problems of disturbed family relationships and methods of influencing parents so that mental disorders in their children may be prevented; next there is a group of five papers concerned with situational and maturational crises in children's personality development and with methods of preventive intervention to

ensure a favourable outcome of such crises; the last three papers deal with the part that schools and teachers can play in promoting mental health. The early chapters indicate that considerable advances are being made in the biological field regarding noxious genetic and organic elements, to which the long list of references to work recently done bears witness. The situation seems to be far less encouraging at present regarding new insights into psychological, social or educational influences upon personality development.

As Dr Caplan sees it, 'much of the practice of preventive psychiatry is conducted through the intermediation of such other helping professionals as teachers, nurses, paediatricians and clergymen'. Though he stresses the manifold difficulties likely to beset this approach, at the same time much is claimed—which may well be challenged by members of allied professions—when he states that, 'The preventive psychiatrist sees himself as an agent of culture change not only in the general community but also among the sister professions'. Finally, Dr Caplan discusses the question of whether there exists 'at present any generally acceptable body of knowledge about pathogenic forces and methods of combating them, upon which we can base an appeal for community action?' He draws a parallel with the history of public health where many of the major advances that had the effect of reducing physical disease throughout the population antedated by many years the science of epidemiology; thus action based upon empirical experiences, plausible hunches and non-scientific hygienic movements succeeded in preventing typhoid and scurvy in Western communities. Moreover, the impetus for the epidemiological studies arose out of the public health programmes rather than vice versa. Similarly, he argues, the problem of mental disorders is so vast and pressing that exploration of possible avenues of prevention seems imperative.

This view is also the starting-point of Dr Caplan's second book. It is addressed primarily to what he calls the 'caretaking agents of the community', namely nurses, doctors, teachers, social workers and clergymen. The material was first presented in a series of lectures to nurses and other groups of professional workers; regrettably there are signs of this throughout the text in the rather colloquial style and the repetition of points which is necessary for listeners but somewhat tedious to the reader. After discussing recent advances in psychoanalytic theory, especially in ego-psychology, an outline is presented of the psychology of pregnancy and of mother-child relationships, including the effects of deprivation and separation. Then the role of the nurse, social worker and family doctor are considered in relation to helping people to find adaptive solutions to life crises; some illustrative examples are given. The final chapter on 'Comprehensive Community Psychiatry' gives one the impression of a blue print, the bare bones of which need to be clothed in explanatory material if it is to be meaningful to the audience for which the main body of the book is intended. Indeed, in its present form the whole book is too conversational for the serious student and too serious for the interested layman; judicious editing for the one audience or the other would greatly enhance its value.

M. KELLMER PRINGLE

*The Psychotherapy Relationship.* By WILLIAM U. SNYDER and B. JUNE SNYDER.  
New York: The Macmillan Co. 1961. Pp. xii+418. £2. 12s. 6d.

This book is a detailed study of the events that took place between one therapist and twenty of his patients, the research focusing on the study of the types of relationship established between the two participants in each therapy, and of the changes in the relationship during the course of therapy. The research method is partly statistical, partly descriptive. An attempt was made to identify and measure the session-to-session changes in affect between therapist and patient, by having both people fill in a detailed questionnaire concerning attitudes towards and feelings about the other person, this being done by both therapist and patient after each therapeutic interview. At various times both patients and therapist completed personality questionnaires and other quantitative tests. This enormous quantity of statistical material was then factorially analysed, the factors which held up statistically being, in the case of the clients, a factor of 'active resistance or hostility', and another of 'passive resistance or withdrawal'. Two negative factors also emerged from the tests completed by the therapist which were 'impatience with the client' and 'anger or irritation with him'. An unidentifiable positive factor was also extracted for the clients and for the therapist. It will come as no surprise to many psychologists that the results from this type of exhaustive statistical approach are incommensurate with the labour involved, and add nothing of value to existing knowledge.



Contrasting with the tediousness of this arid material, it is during the verbatim descriptions of the therapeutic process that the book really comes to life. The reader is given the privilege of overhearing some of the interviews, and anybody interested in exactly what happens in one particular type of eclectic non-psychoanalytic psychotherapy will find much of value from these sections of the book.

Although viewed as an attempt scientifically to study the relationships between a therapist and his patient, the overall result is disappointing, the publication of the book marks an important stage in the current growing awareness that the relationship between therapist and patient is an essential core of the psychotherapeutic process.

J. L. BOREHAM

*Though this be Madness: A Study in Psychotic Art.* By GEORG SCHMIDT, HANS STECK and ALFRED BADER, with a Foreword by JEAN COCTEAU. London: Thames and Hudson, 1961. Pp. 114. £2. 2s. 0d.

Apparently this book is a translation, because it is stated that the original editions of it were published under the title *Insania Pingens*, with the support of CIBA Ltd., Basle. The book consists of several sections. Jean Cocteau contributes a foreword entitled 'Minor Masters of Madness'. Georg Schmidt discusses the question, 'What has the art of psychotics to do with art as such?' Hans Steck deals with 'Primitive Mentality and Magical Thought in Schizophrenics'. Alfred Bader's section is entitled, 'The Pictorial Work of Psychotics—a Mirror of the Human Soul'. These chapters are followed by sections about three psychotics whose pictures are illustrated in the book: Aloyse, Jules and John. There are notes to the text, a bibliography and a table of illustrations, 107 in all, almost one to every page.

Cocteau's theme is that one must be a little mad to be a genius, even if it is not apparent, as it was in some artists like Van Gogh. He thinks genius frightens us because it is disobedience and expresses the spirit of non-conformity. This is a doubtful theme, unless we interpret the idea of madness—or schizophrenia, as he calls it—very liberally.

Schmidt, in his part, shows that the horizons of our artistic appreciations have been very much widened in recent times. We can see artistic interest in the popular work of the so-called *peintres naïfs* of the nineteenth and twentieth centuries, the drawings of children, and the art of the mentally abnormal. However, the art of children differs fundamentally, in his opinion, from that of adults, and one has to develop to certain maturity in order to approach the possibility of real artistic achievement. In a similar way, the art of the psychotic is limited. It is not intended for the public and not even for the patient's own companions. It is private art, and insanity does not trigger off artistic genius, except in very rare cases. This did not happen even in Van Gogh, whose art was greater at the times when he was less disturbed. The fact that we have learned to see artistic interest and merit in the paintings of psychotics does not show that their work is the output of genius, unless they have artistic genius unimpaired by mental illness or released indirectly by it.

Steck shows how in the mentality of psychotics there are interesting resemblances to the mentality of so-called primitive peoples. In his mental regression the psychotic falls back to levels of magical, mystical and non-logical thinking, as described by such writers as Levy-Bruhl and Frazer. Many examples are given of this, illuminated by discussions of the paintings in the book.

Bader explains the peculiarities of schizophrenic art with numerous and very cleverly chosen illustrations. He shows how psychotic art penetrates into the unconscious, or rather reveals it, and has a special appeal due to the universal nature of its primitive imagery. Objects interpenetrate each other, undergo strange and unreal transformations and colour may be unrelated to external reality. Although the aesthetic value, such as it is, may be smothered by the pathological qualities of the imagery, psychotic art is often very attractive, and may be reminiscent of the work of great artists, like Klee or Picasso, who skilfully exploit similar subject matter and techniques of expression, without themselves being psychotic.

The paintings illustrated in this book are fascinating, even if not often aesthetically pleasing to a high degree, and are extremely well worth having as examples of psychotic art. Many of them are beautifully reproduced in colour, and it is to be hoped that other collections of the works of schizophrenic or other patients will be published with equally wise and balanced commentaries. It is not an unreasonable suggestion that some art galleries might make collections of the works of psychotics, and this might help to reduce the very understandable fear

which normal artists have of psychotic work. There is no real danger that the art of the abnormal will ever undermine good art, however exciting it may seem. The real contrast is not between psychotic and normal art, but between good art and bad, and few psychotics would succeed in producing exceptional works of art because they have neither the training nor the technique, although some of them exhibit talent and inspiration of unusual quality. Their art is interesting for other reasons, as this book shows very clearly.

R. W. PICKFORD

*Contemporary European Psychiatry.* Edited by L. BELLAK. London: Evergreen Books. 1961. Pp. xxvi+372. 25s.

*Transactional Analysis in Psychotherapy.* By E. BERNER. London: Evergreen Books. 1961. Pp. 270. 25s.

Dr Bellak has written an amusing introduction to his book, in which he compares European and American psychiatry and makes a few well-deserved criticisms of the organization of research in the United States. This is followed by a series of excellent contributions by distinguished psychiatrists on the state of their speciality in France, Great Britain, Italy, Scandinavia, Switzerland and the Soviet Union. However, the section on German psychiatry, which played a leading role in European psychiatry from 1850 to 1933, is virtually unreadable. It has been written by two distinguished Viennese psychiatrists and translated into English by someone who is unacquainted with German psychiatry and the English tongue. Thus Kleist's term 'phasophrenie' is rendered into English as 'phasic schizophrenia', whereas Kleist used this word to designate *all* phasic functional psychoses, i.e. manic depressive and schizo-affective psychoses. The interesting ideas of Carl Schneider and Klaus Conrad, again, are presented in an incomprehensible way.

Dr Berner's book is a presentation of the personal views of a successful psychotherapist and is written in racy, neo-psychoanalytic American, which amuses, startles and nauseates by turns. He suggests that there are three ego states which are expressed in the individual's behaviour. These states are the parent, the adult and the child. The patient during group treatment learns how these ego states affect his behaviour and also what game he is playing in life.

Transactional analysis consists, therefore, of a group analysis of the ego states and games of the members of the group. This analysis can be carried even further and one of the 'New Frontiers' of this therapy is regression analysis in which the patient is told to assume that the therapist is five years old and the patient is less than eight years. The therapist then plays the part of a child who does not understand complicated statements!

It is easy to be scornful of this book, because the author's brash statements are irritating, but devotion to his patients and his obvious enthusiasm for his work are apparent despite everything. Once again one is forced to the conclusion that successful psychotherapy depends on the personality of the therapist rather than on his theoretical system.

F. J. FISH

*Der Politische Charakter.* By FELIX SCHERKE. Stuttgart: Hirzel. 1961. Pp. 115. DM 12.60.

This is a short book, without index, which might almost be used as the perfect illustration for the widespread belief that German psychology is the Rip Van Winkle of our science. There is no awareness in it of the need for facts, for verification, indeed for anything but philosophical speculation. There is no recognition of the tremendous amount of empirical work done by social psychologists in English-speaking countries; indeed the references in this book are evenly divided between those to Fritz Kunkel: *Grundzüge der politischen Charakterkunde* (1934), and to a variety of other German writers who, like Kunkel, wrote their books and articles at a time when Hitler had not even become Chancellor. The reader will not be surprised that psychology is defined as 'the science (Wissenschaft) of the conditions, events, causal factors and appearances of the conscious and unconscious soul and mind-processes of individual people and of groups'; no mention here of such a vulgar thing as behaviour, measurement, or experiment. Readers who are interested in seeing how little the writings of a psychologist in Germany differ from those of a third-rate leader writer in this country may be recommended to glance at this book; those interested in the application of psychology to politics cannot be advised to do so.

H. J. EYSENCK

*Anleitung zur gestuften Aktivhypnose.* By D. LANGEN. Stuttgart: Thieme. 1961. Pp. iv + 48. DM 4.80.

This small book illustrates a movement which has been gaining impetus in the past 10 years, that which advocates a more dynamic psychotherapy through the production of a hypnoidal state in the patient.

After a historical introduction the author concentrates on a description of the practice of 'graduated active hypnosis' which was elaborated by E. Kretschmer on the basis of J. H. Schultz's *Autogenic Training*.

There is a thorough investigation of the patient's history and character until the conflict is brought into the open. This is called 'conflict-centred therapy'. The situation is then analysed. Resort is had to key slogans to bring about a change of outlook so that the symptoms retreat into the background. The patient's assets are mobilised, especially his drives and his intellect. The patients not infrequently remain in a state of passive concentration for as long as an hour during which time they may experience cataleptic or anaesthetic symptoms. One may add that the changes which occur in the brain and autonomic nervous system are controlled in many clinics by polygraphic recordings of EEG, EKG, respiration, skin resistance, myography, etc., and many interesting discoveries have been made in the field of neurophysiology.

The treatment is particularly successful in treating intractable pain and results are comparable to those of leucotomy. At the end of the course of treatment the patient is left with the ability to relax himself to a maximal degree when desired and to give himself suggestions when in the state of passive concentration. The author might have added that in Germany in the last few years, some 10,000 patients have been treated by these or similar methods. The book is recommended to anyone anxious to keep abreast of modern trends in psychotherapy. It has a full and international bibliography.

A. SPENCER PATERSON

*Contemporary Issues in Thematic Apperceptive Methods.* Edited by JEROME KAGAN and GERALD S. LESSER. Springfield, Ill.: Thomas. 1961. Pp. xiv + 328. 96s.

The book would be more correctly called 'Proceedings of a conference on apperceptive methods (mainly T.A.T.)'. The conference, which took place in America (where? when?), was motivated by a desire to determine a rationale for apperceptive methods supported by experimental evidence. There were sixteen participants, all well known in that particular field. The publication covers the papers read by eight of them and each paper is followed by discussion opened by one of the other eight.

Since Sanford's pioneer work attention has increasingly been directed towards the forces which influence the manifest content of T.A.T. stories. The present position is reflected in the papers which are broadly concerned with defence mechanisms affecting the production of the stories. Much pertinent research appears to be in progress in the departments headed by the contributors and one paper deals with a nation-wide survey on social motivation. The reviewer was attracted by the two more general papers: one analyses phantasy, daydreaming, reverie, etc., the other, starting from Aristotelian logic, discusses patterns of thinking.

Each lecture taken in isolation has something to offer. The cumulative result is, however, intellectual indigestion. For research and references the compilation has its uses but, had they been rigorously edited and judiciously distributed over the technical journals, the papers might have been more effective.

A. KALDEGG

*Sensory Communication.* Edited by WALTER A. ROSENBLITH. New York: Wiley. 1961. Pp. xiv + 844. 128s.

This gives the substance of a Conference held at the Massachusetts Institute of Technology in July 1959: 38 chapters of original contributions, a chapter of comments by some of the contributors, and a final statement of the present position by the Editor. Almost half the contributors come from outside North America, four (Barlow, Cherry, MacKay and Rushton) from this country. Over half the chapters are straight neurophysiology, a quarter cover psychophysics and perception, and most of the remainder are sensory physiology.



For the Part II and research student in Psychology the book is to be recommended principally for its good coverage of neurophysiology, since it collects together recent work scattered fairly widely throughout the literature, much of it not available in the standard American journals, and some not printed in English. The coverage of the remaining topics is patchy. Thus on the measurement of sensory intensity, the field with which your reviewer is most familiar, a rather optimistic picture is presented of the results which have been and can be achieved, without reference to many of the intractable methodological problems which should still be facing the research worker. However, despite these limitations, the book will make a useful addition to our Unit library.

E. C. POULTON

*Gifted Children.* By M. F. FREEHILL. New York: The Macmillan Co. 1961. Pp. ix + 412. 38s. 6d.

Since the Sputnik of 1957 there has been a growing concern in the U.S.A over the products of its educational system. There have been strong demands for the early identification of talented children, and for an adequate educational programme for them. The view that the segregation of bright children is undemocratic—a view which might well be primarily responsible for the national denigration and suspicion of 'eggheads'—is more and more being questioned. The 'academic lock-step' based on the credit point system and a denial of accelerated school progress has come in for some critical examination. This book reflects the trend. But it is no clarion call to revolution. It lacks the incisiveness of a Terman or a Barzun: it is the text of the instructor who is already convinced and needs guidance on 'how to do it'. But in spite of the cautious approach and the leaning-over-backwards, Mr Freehill gives a solid survey of what is being done in some states and in some school systems, gives sound advice on methods of 'identification', methods of teaching, the characteristics of highly intelligent children, the differences between 'enrichment' and 'acceleration' and so on. He draws no evidence from other national systems (except for a brief flirt with Canada) but shows clearly that the movement in the States is as yet a pretty small one.

STEPHEN WISEMAN

*Professional Satisfaction among Swedish Bank Employees.* By UNO REMITZ. Copenhagen: Munksgaard. 1960. Pp. i + 422. Dan. kr. 86.00.

The problem of satisfaction at work has aroused a good deal of interest over recent years, and has spread from an interest in manual to the 'white collar' workers. Dr Remitz's study is one of the most recent and most ambitious. The book can be considered from two viewpoints: that of the way in which the investigation was conceived, planned and carried out, and that of the very sophisticated statistical methods which were used in analysing the data.

The survey took place on a sample of some 110 Swedish bank employees and was possible only through the active co-operation of their trade union. The assessment of satisfaction was by means of questionnaires of substantial dimensions. The usual areas of working experience were tapped: the work itself, colleagues, relations with supervisors, management policy, welfare policy, etc. In addition to giving a measure of satisfaction for each item, a measure was given of its importance to the respondent. It is over the treatment of these two assessments that the statistically minded will be able to enjoy their argument and disagreement, and will question the validity of the findings. These can be summarized:

The level of satisfaction is a primary psychic function in its own right unrelated to sex, intelligence, extrovert-introvert personality or neurotic disturbances. A variety of character differences between the more satisfied and the less satisfied can be shown. External factors which most determine the level of satisfaction are: position in the bank, salary (with higher than 'minimum' a strong factor for satisfaction), education and health.

The book is provocative, it invites examination of the hypotheses that are advanced, and its methodology will stimulate discussion. The reviewer cannot but marvel that a Finnish author, reporting a survey in Sweden, and publishing in Copenhagen, can produce a work in such polished English, with clear charts, diagrams and tables, and generally so excellently finished a production.

H. G. MAULE

*Journal of Psychiatric Research*. London and New York: Pergamon. Price £7 per annum (£3. 10s. for individuals). Vol. 1, No. 1 (October 1961).

Here is a new journal. The aspiring author already has a wide choice—*Archives of General Psychiatry*, *American Journal of Psychiatry*, *Journal of Nervous and Mental Diseases*, *Journal of Mental Science*—quite apart from fringe journals and weeklies such as the *Lancet*.

The first issue contains no editorial notice of the reasons for its existence, nor a statement of policy. Perhaps the publishers decided on a sister for their *Journal of Psychosomatic Research*, which recently held out tempting offers of quick publication to prospective authors. The new Editor-in-Chief, Seymour S. Kety, has hitherto devoted himself to a distinguished career in chemical physiology, and although the prospective journal buyer will see distinguished British names among the editorial board, only two among the seven are those of psychiatrists.

Some pertinent criticisms of schizophrenic twin research techniques are contributed by Rosenthal. The other seven papers seem to have been permitted more space than their insubstantial positive results merited. Forty of the 100 pages are taken up by the Editor and his immediate colleagues. Their studies of the effects of injection of epinephrine into controls and schizophrenics were obviously conducted with great care, but one would have been surprised had a different journal allowed 4000 words for their third paper, reporting only expected negative results. Much the same could be said of certain other contributions. One hopes that there may be forthcoming enough psychiatric research reports, both succinct and of good quality, to satisfy all the competing journals.

IAN OSWALD

*Problems of Psychology (Voprosy Psikhologii)*. Edited by N. O'CONNOR. Oxford and London: Pergamon. 1961. Nos. 1 and 2. Pp. 160. Private subscribers, £5 per annum; Libraries, £15 annually.

This long-awaited journal, the publication of which is sponsored and assisted by the British Psychological Society, has now been launched. This double number contains 14 original papers translated in full from the Russian journal *Voprosy Psikhologii* as well as abstracts of 11 more. The papers have been selected to give a picture of the best work being done currently in Moscow, Leningrad, Tbilisi and a few other centres of psychological research in the Soviet Union. The fields of major interest to Soviet investigators are well represented here in papers by Leont'ev on abilities, by Zinchenko and Lomov on hand and eye movements in perception, by Gan'kova on the relations between action, image and speech, by Schwartz on conditioned verbal reflexes, by Luria on memory, by Sokolov on a probability model of perception and other articles on a similar range of topics using children as subjects. The scientific level of the work is high even if the techniques strike us as rather old-fashioned. This is shown in several ways: the minimal use of statistics, the study of a single case (Luria's 30 years' study of a man with an extraordinary memory, amounting to almost total recall), the central character of the problems of perception, the relative absence of jargon, the study of individual learning in a real-life situation with a 'commonsense' level of analysis, and so on. Perhaps the least successful paper is that by Leont'ev on the 'formation of abilities', the only generalizing article in the selection. The translation is on a really deplorable level; without consulting the original it is impossible to understand the principles Leont'ev is developing. Proof-reading and editorial work generally seem to be based on a policy of non-intervention. The binding is slipshod so that even before reading the journal had already come apart. Pergamon Press can do a lot better than this!

Supplementing the quarterly publication of this journal there is a service whereby translations of any article abstracted can be supplied at a reasonable price. There would seem to be no longer any good reason why there should be continued ignorance of Russian work on psychological problems.

JOHN MCLEISH

## EDITORIAL

### THE *B.J.P.* IN JUBILEE YEAR

Reference was made in an earlier issue (Vol. 52, Part 4, p. 469) to the Diamond Jubilee of the British Psychological Society. The present issue is offered as a 'Jubilee number'.

In place of the promised article on Psychology during the Society's early years we are printing a version of Prof. Pear's after-dinner talk at the 1961 Christmas Conference. We felt that this was more appropriate to the occasion and more likely to be of interest to readers than what would probably have emerged as a somewhat academic exercise.

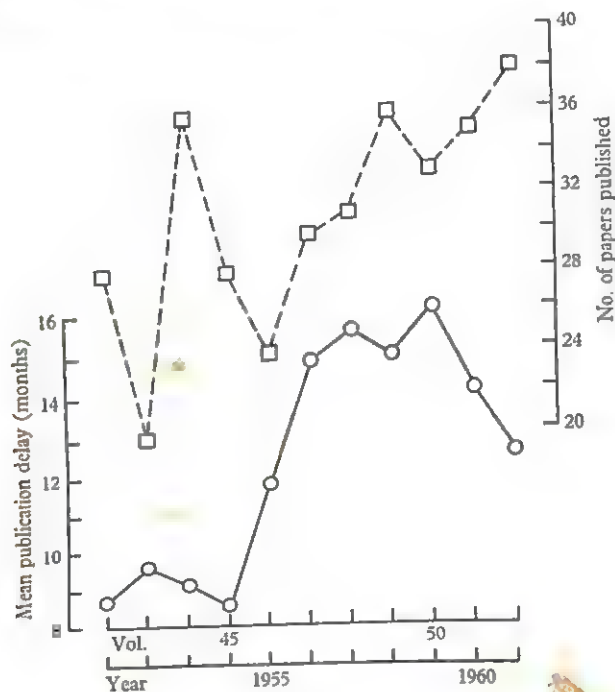


Fig. 1

The keen-eyed reader will have noticed that the entry '*Manuscript received...*' is missing not only from Prof. Pear's paper but also from Sir Cyril Burt's. Prof. Burt offered his paper just when plans for this issue were maturing, and we felt that no offer could have been more opportune. We have therefore given this paper priority of publication as a second contribution to the Jubilee number.

Mention of date of receipt brings us to our third Jubilee offering. It is seldom that the Editor can directly address the readers of the *Journal*, and we felt that we might take this opportunity of answering the question perhaps most frequently asked: 'How long shall I have to wait for my paper to appear?' Strictly speaking there is no such thing as immediate publication in a scientific journal: the irreducible minimum is three to six months. The Society's Council is anxious that the publication delay



should be as short as possible, preferably not more than nine months, and with this end in view increased page allocations have been made. Some idea of the extent to which the goal is being reached may be obtained from Fig. 1, which shows the average publication delay, in months (full line), together with the number of papers published (broken line), for successive volumes starting from Vol. 42 (1951), the point at which the Volume was adjusted to coincide with the calendar year. Various considerations prevent data of this kind from being entirely unambiguous, but random sampling among earlier volumes has suggested that the position has never been much better than about midway between the present state of affairs and the nine-month target.

Delving in the archives has brought to light a wealth of facts and figures. We plan to prepare an article on 'Publication trends in the *British Journal of Psychology*', to appear in the *Bulletin* during 1964, which will be Jubilee Year for the *Journal*, as distinct from the Society.

### CORRIGENDA

A number of misprints occurred in the list of References appended to the article by H. M. B. HURWITZ in Vol. 53, Part 2, 167-173. The entries should read as follows:

HURWITZ, H. M. B. (1958). A source of error in estimating the number of reinforcements in a lever-pressing apparatus. *J. Exp. Anal. Behavior*, 2, 149-52.

HURWITZ, H. M. B. (1962). The effect of a discriminative stimulus on behaviour under a fixed ratio schedule of reinforcement. (In preparation).

HURWITZ, H. M. B., BRENER, J. & JONES, B. (1961). Some experiments on behavioural stability. *Bull. Brit. Psychol. Soc.* 43, 2A.

Also in the reference to MECHNER, F. M. (1958), for 'ration' read 'ratio'.

In the same issue (article by R. T. GREEN), p. 111, line 8, for 'significantly' read 'sufficiently'.

## PERSONALITIES IN THE EARLY DAYS OF THE BRITISH PSYCHOLOGICAL SOCIETY

By T. H. PEAR

*Emeritus Professor of Psychology in the University of Manchester*

To disclose the circumstances which produced this article may explain its unusual form as a contribution to this *Journal*. The editor asked me to write up a light-hearted speech made at a dinner held in London on 15 December 1961, the year of the Society's Diamond Jubilee. Impulsively I assented; at leisure I nearly repent. I had hinted at the personalities of early members of the Society as they appeared to a young recruit. The Society's ideals and the outstanding contributions of its founders are discussed in Prof. L. S. Hearnshaw's 'Sixty Years of Psychology' (*Bulletin of the British Psychological Society*, no. 46, January 1962, pp. 2-10). Photographs of the psychologists mentioned below are in the Society's archives.

'Geography', wrote Nicholas Bentley, 'is about maps; biography is about chaps.' People who are not well informed about psychologists' fashions are apt to suppose that psychology, too, must be about chaps: a day spent among the current journals might shake this belief. Yet the British Psychological Society was founded and is run by chaps. Their peculiarities and special interests will always influence its development. May I therefore speak of some of the pioneers?

If one lived in the north, to attend an afternoon meeting of the Society in London it was necessary to rise early. So it might be daunting to find at King's College at 2.30 p.m. no signs of life except an aloof cat and a bonhomous hall porter. He would assure you that the meeting was certainly 'on', but 'the gentlemen are never very punctual, sir; *you* understand'. Sure enough, about 2.35 p.m. a small chorus of dons could be heard advancing up the corridor (hearing aids were still below the horizon) and up would stroll the in-group, to be augmented by stragglers, until at 2.45 p.m. there might be thirty in the meeting-room. Not many, but representing philosophy, education, neurology, psychiatry and anthropology, and perhaps a third of the audience were already known in the English- and German-speaking worlds of knowledge. One might expect to meet a goodly number of the following: S. Alexander, F. A. P. Aveling, P. B. Ballard, W. Brown, E. Bullough, Wildon Carr, Beatrice Edgell, Bernard Hart, H. Head, C. W. Kimmins, T. W. Mitchell, C. S. Myers, W. H. R. Rivers, A. F. Shand, C. E. Spearman, G. F. Stout. And—for in those days the Society was a very British one—we confidently expected, and got TEA. At these breaks older members took pleasure in talking to the newcomers. I remember the pleasure of hearing Sully tell me that, now he had read Freud, he wished he were young enough to re-think the whole problem of dreams.

The early Society had no sections, so the papers read related to a very broad range of subjects, and were criticized by members with wide interests, not only in psychology and philosophy. Pure psychologists were heard eagerly by iconoclastic young psychiatrists; a paper on intelligence tests might be regarded by an anthropologist as rather parochial in outlook. At times, perhaps, it might be felt that mathematically gifted psychologists could hardly be expected to minimize the importance to any good life of rapidly seeing mathematical relationships, even if such

flights of intelligence were not highly correlated with ability to see a joke. A paper on time estimation might be praised by a philosopher—and so on. To these toughs, uncommitted to any theory, let alone any policy, it would have been hard to justify the use of a particular mental test on the ground that it had been standardized, or that it obviated any embarrassing face-to-face encounter of tester and testee. No member represented a group or circle, though their existence on the continent was noted. That any school of thought in science might turn into a power-group might have seemed unthinkable then.

The average member did not attend primarily to learn more about his subsection of knowledge, but to enjoy the *table d'hôte*. The menu might offer McDougall or Bullough on colour, Myers on primitive music, Spearman on intelligence, Rivers or Head on the cutaneous senses. It was a mosaic, even a kaleidoscope, but with hints of incipient *Gestalten*. Occasionally it was delightful to see a jungle of thought being opened up: seldom did one feel that one's mind was merely being taken for a suburban ride on tramlines, or a monorail. Democratic expression of opinion was the rule: if a member hadn't grasped a point he was apt to say so, even if some hearers doubted whether he had tried with all his might.

Symposia were held, in the literal as well as the metaphorical sense. Notices of London meetings indicated that if names were sent in advance to Mr Shand, a dinner would be arranged. After a few visits, a senior member prodded me to apply, but at the meeting I heard with sadness that the quota had not been reached. A surprise was in store: while a speaker was in full spate, Shand flipped across to me a note with the scrawl 'Commercio, 7.30'. There I found a lively party and a pre-war Soho spread. Quietly someone explained to me that if Shand ever got a name he didn't like he just washed out the Society's dinner and made up his own list. Shand's hedonism was shown in many ways. He would describe the Sentiments, especially Love (giving the auditory equivalent of capital letters) like an epicure recommending rare wines. Just after World War I, when money was tight for a young married lecturer, he exhorted me to be very sceptical of any Manchester vintner who pretended that he sold only green Chartreuse. 'Stand over him: he'll bring out some yellow!'

Here are a few more impressions: first of the men to whom I owe so much, Myers and Spearman. Their different *Weltanschauungen* coloured their methods of imparting knowledge. By Myers, scientist, medical man, musician, traveller, anthropologist, any subject he discussed, in or out of the lecture room, was likely to be illuminated from some entirely unexpected source. (You remember that when the Winged Victory is floodlighted, she moves forward surprisingly because she is being successively lighted from more than one direction.) And this, combined with Myers's going rather sticky for a student who expected his lecture notes to supply packaged answers to all conceivable examination questions—or what were lectures for? Spearman, in casual conversation and in lecturing, talked decisively, with the special brand of courtesy characterizing his first profession—that of a regular army officer. At times, indeed, I felt that his tone was that of an army instructor rather than of a lecturer; that for him the mind, like a rifle, had parts, the names and functions of which must be learned before one proceeded to the noegenetic principles—Army



Council Instructions on the ways in which the mind ought to work. Yet he did not disapprove of introspection by other psychologists, for he sent on my behalf a letter of introduction to Külpe.

That I was privileged to benefit from such contrasted approaches to psychology was due to Spearman's generosity. When Myers left England to winter in Egypt, Spearman warmly invited me to join his Honours school, and gave me every possible facility and encouragement. A third introduction to psychology was to come my way, for Manchester University allowed me to join the famous set-up at Würzburg. Külpe would expound a theory with enthusiasm, giving it (in today's elegant phrase) all he'd got, then, with a twinkle and a forefinger laid alongside his nose, he might comment '*Aber, freilich!*'. And the bottom brick from the edifice might be deftly removed.

Curiously perhaps, my love for the psychology of individual differences seems to owe little directly to these early teachings, except that to any student of psychology in the Strand or Gower Street Galton was a hero. (Cf. Sir Cyril Burt's '*Galton's Contributions to Psychology*', *Bulletin of the British Psychological Society*, no. 45, September 1961, pp. 10-21.) It impressed me then—and has done so ever since—that differences of predominant imagery are a source of delight and wonder to a few, and of potential embarrassment to many psychologists, though the latter have been assiduous in constructing ingenious (and dull) by-passes.

In the early days the names of Head and Rivers were inevitably linked with their pioneer nerve-section experiment. Later, during much work with Rivers at Maghull Military Hospital, Liverpool, I heard a great deal about how his mind, and Head's, worked. So far as I know, in the years after the first war, only Robert Graves and I were interested in the spectacular differences between the mental equipment of these men whose co-operation was so smooth and perfect. Head was a vivid visualizer, Rivers an habitual verbalizer. Rivers used to regard Head's mental camera as a vestigial mechanism of 'lowly organization' out of which he ought to have grown after childhood. Rivers once declined to write down 'as a contribution to science' Head's excited declaration during an experiment that the sensation of cold returning after long absence 'entered like a light into darkness'. But when Rivers began his autoanalysis of dreams, he became impressed by the importance and clarity of his visual images. (From his account of them, I suspected that they were rather smudgy and dim.)

Once he asked me how, when lecturing from skeleton notes, I could quickly produce and elaborate several concrete examples of a general statement. When I replied that the examples often came first, as the basis for the formulation, he was very surprised, as his mind worked, and slowly, from abstract to concrete. In his *Poetic Unreason*, Graves has analysed the meaning of one of Rivers's sentences (in *Instinct and the Unconscious*) which is packed with visual imagery, and concluded that this use was far from successful.

Rivers described to me his method of working when at college. For hours in the early morning he would lie in bed planning an article—in words. He would meditate over a solitary breakfast and for the rest of the morning would type deliberately—even with one finger—the words which he had imaged, hours before, presumably in auditomotor terms.

From 1910 onwards, following Freud's visit to Clark University, and the subsequent publication of the famous special number of the *American Journal of Psychology*, there was daily talk of him among the younger English psychologists, but I doubt if the new ideas disturbed the oldsters much. Certainly not Ward. And during a symposium of psychologists and philosophers at Durham (those were pre-apartheid days) on The Role of Repression in Forgetting, McDougall, in the chair, seemed to more hearers than Ernest Jones to have done insufficient homework on Freud's general concept of memory. Jones suggested that perhaps, in one sense, all forgetting in the healthy mind might be the result of repression of the irrelevant, cognitively or affectively. This really moved Stout. He stumped down the lecture room stairs, and towering above the seated Jones, demanded 'Did I hear you say that ALL forgetting is due to—an ominous wad of silence—repression?' Politely Jones began, 'In a sense, yes. . .'. 'Ugh!' said Stout, and swept to his seat. I have heard lengthier criticisms of Freud.

As my arrival in the new post at Manchester resulted from a simultaneous push from Myers and a pull from Alexander, I came to know this legendary giant well. (Most of the classical stories about him are apocryphal, but the true ones are funnier.) He was glad to hand over to me the teaching in psychology as, an early cosmonaut, he was pondering, *ambulando* and on his beloved bicycle, Space, Time and Deity. I think that some developments of psychology at that time would have been hard to fit in with his elegant theories. Like simple behaviourists of that era—but simple, S.A. was not—he looked snags straight in the face and passed on. Towards dreams, the structure of which, to my mind at least, challenged his whole system of realism, his attitude was as ribald as that of a run-of-the-mill physicist, and perhaps for the same reason—to preserve mental comfort. He maintained, at least in conversation, that he had heard of no dream event which contradicted any physical law. As for Freud's concepts of symbolism, condensation, and regard for presentability, did Alexander repress doubts they might suggest? When, in Alexander and Stout, Australian sturdiness clashed with Tyneside toughness, it was fun to be on the side-lines. Alexander once hauled Stout on to a platform, announcing: 'Last week in London Mr Stout didn't seem to hear some of my objections to his views on sense data. So that he shall hear perfectly tonight, I now give him a copy of the paper I am about to read.'

In 1923, the Seventh International Congress of Psychology was held in Oxford. Though Morton Prince probably knew Stephen Leacock's eulogy of an Oxford education ('being smoked over, and sleeping in the ivy, but the ivy's influence is the more important') he did not wholeheartedly endorse this Englishman's nostalgic opinion. I was told that on arriving at the ancient college which had the honour of housing him, he naturally informed the scout that he would like a bath. 'Yes, sir, certainly' was the optimistic reply. 'Go through the first and the second quad, turn left by the big tree, bear round' (or was it climb over?) 'the rockery. . .'. Prince just remarked, tight-lipped, that the Oxford boys must have found the trenches in France quite luxurious. Pierre Janet sat through an Englishman's rather wide-meshed address till, meandering into a delta of vagueness, it dried up. Janet did not join in the faint applause, remarking briskly to Alexander, 'But it could be cured!'

That inward eye which is the bliss of solitude still vouchsafes me a delightful

memory of four candle-lit faces in a college hall. Conversing happily were Arthur James Balfour, Henri Bergson, Samuel Alexander and William Archibald Spooner, the voice who launched a thousand slips. He had finished his address by purring 'But I will not go on, or I might be tempted to say one of those things that people say I say...'.

Who ever spoke of social psychology at those meetings? Almost all laboratory experiments attempted to answer questions framed with no social implications. Usually they employed persons, as often as not undergraduates, who, unless defective in some one sense organ, were assumed to typify mankind in general. 'Other things being equal', prefacing a conclusion drawn from an experiment, was sometimes a face-saving clause, obscuring the fact that other things usually weren't. And in textbooks of general psychology, the chapter headings were apt to be Instincts, Emotions, Volition, Language, etc. David Katz remarked to me, about 1935, that psychology textbooks written by Germans, French, Americans, English, even Japanese, might have almost similar general headings. There were, of course, signs that the study of individual differences might soon be very important: Binet's account of Marguerite and Armande, Stern's of the conflicting testimony of witnesses of the same event, the story of Helen Keller, the descriptions of lightning calculators and chess players. But *social psychology*?

It was not easy to distinguish the aims of early social psychology from those of sociology. As Hadley Cantril pointed out in his article 'The Social Psychology of Everyday Life', it seemed then to be the social psychologists who loved grey generalizations and shied away from concrete examples, which—it seems hard to believe this today—were supplied by some sociologists. (There are still prominent students of society who never seem really to have liked or approved the existence of actual persons.) McDougall's *Introduction to Social Psychology* had not appeared yet. Mental tests and factors were still being described and discussed as interesting developments of theory. Who could have foreseen the day when, at least according to the Press, the phrase 'eleven plus' could be used to a child in the minatory tone with which sadistic nurses pronounced 'Boney will get you!', when Martello towers were rising on the south coast?

I doubt if in the early days of the Society many psychologists thought of themselves as particularly useful, in the narrow sense of that word. Some had heard the Cambridge mathematicians' toast: 'And may this discovery never be of any use!' The fashionable assertion: 'I study this because it interests *me*', may have seemed reasonable enough. In those days it occurred only to livelier minds that psychologists might have something to say concerning the mentally deficient, the blind and deaf. But for many, 1914 meant the end of non-useful speculation. The utilization of psychology in World War II has been fully chronicled, but from 1914 to 1918 there was little time to write things up, and with the coming of peace many teams were scattered for ever.

It has taken me years to recover an abiding and intense interest in any psychological problem which seems unlikely to yield a socially useful result. My interest in human social psychology (I have never had a satisfying conversation with a rat) results from a liking for persons, particularly those between adolescence and maturity—are they the forgotten men and women of psychology? For this interest some



reasons are not far to seek. As soldiers these people had been studied—in some ways they were a captive collection—but not as civilians anxious to begin a new life. Before 1919, for four or five years I had lived with, and observed, in England and Germany, hundreds of psychotic and psychoneurotic persons—to me the arbitrary Procrustean lopping off of 'psycho' still seems unjustifiable for scientific reasons—and even off duty the few acquaintances I made were living in abnormal war conditions (I had no need to read imaginary accounts in the spate of post-war novels). Emerging gradually into 'peace', I formed the impression that psychologists would probably focus their attention on infants, children, subnormal and disturbed minds, delinquents and criminals, and that, while stimulated by Freudian theories, there would be much talk of envy, hatred, malice and uncharitableness, little attention would be paid to friendly, laughter-loving people who got on with parents or children and seldom went to psychiatrists or to prison, and whose antics were not sufficiently bizarre to attract the attention of Somerset Maugham or Noel Coward.

McDougall revived an old interest in London life, and Trotter's *Instincts of the Herd in Peace and War* introduced me to concepts of the sensitive and resistive mind, related to Jung's introverts and extraverts. Trotter's picture of the university don, protected from the rigours of ordinary—if not of college—life, so that he might think valuable thoughts about life in general, I still find suggestive. Such reading encouraged a critical approach to any account of behaviour which ignored the social moulding, especially through speech, of all human actions.

A few psychologists were grasping the social-psychological undertones in Frederick W. Taylor's bull-at-a-gate bash at Motion Study, and the resistance of some economists and trade union leaders to 'Scientific' Management. (Science, what crimes are committed in thy name!) But the popular interest in rumour, class-feeling, taste, fashions and their deliberate manipulation, conflicts and wars, national prejudices—subjects which the man in the street might be forgiven for assuming to be the distinctive study of psychologists—was not shared inside the universities. Interesting subjects were not yet academically respectable.

Sometimes I am inclined to wonder (perhaps disloyally, but the subject of scientists' proper loyalties is tricky and usually avoided) how many early members of the Society ever thought about social-psychological issues, in the sense in which that term is used today. Stout wrote about intersubjective intercourse; anthropologists described the manners of primitive tribes, but how many saw that Shaw's *Pygmalion* underlined the social implications of speech and clothing in our own society?

On my way to the dinner, hundreds of Christmas gifts competed for attention in the shop windows. One was an expensive box containing multi-coloured sections of Swiss cheeses, insulated from each other by tinfoil. Possibly it is marketed by a Swiss friend. On his hospitable table I have never seen these packets, but instead, a wedge of Gruyère, creamier than anything sealed in these gaudy wrappings. Perhaps the sectionizing of any large Society is inevitable, but need the coverings be so unsusceptible to osmosis? Must the proper study of mankind involve so much processing, and perhaps, loss of flavour?

# THE CONCEPT OF CONSCIOUSNESS

By CYRIL BURT

*University College London*

The object of the following paper is to criticize the current practice which takes behaviour rather than consciousness to be the defining characteristic of psychology. It is argued that consciousness must be regarded as a generic term covering (i) certain specific relations (e.g. intuitive awareness), and (ii) certain specific contents (e.g. sense-data, mental images, feelings, etc.). Both relations and contents are directly given; both are unique phenomena of sufficient importance, practically as well as theoretically, to deserve systematic study in their own right. The methods available for studying them necessarily include introspection; and this, it is contended, is as valid as any other mode of observation. Detailed instances are given, showing how the neglect of introspective procedures leads to descriptions of human behaviour which are not only incomplete but frequently misleading.

## I. CURRENT CRITICISMS OF THE CONCEPT

'The time has come when psychology must discard all reference to consciousness, and need no longer delude itself into making mental states the object of observation; its sole task is the prediction and control of *behaviour*; and introspection can form no part of its method.' Nearly half a century has passed since Watson (1913) thus proclaimed his manifesto. Today, apart from a few minor reservations, the vast majority of psychologists, both in this country and in America, still follow his lead. The result, as a cynical onlooker might be tempted to say, is that psychology, having first bargained away its soul and then gone out of its mind, seems now, as it faces an untimely end, to have lost all consciousness.

In the hope of rescuing it from the threat of extinction, I want, if I may, to enter a plea for the reinstatement of consciousness as a useful and necessary concept, and at the same time try to rehabilitate introspection as a valid scientific procedure. The natural starting-point will be a brief examination of the arguments commonly advanced in support of the opposite doctrine. They have been re-stated in up-to-date form by several American writers (see Mandler & Kessen, 1959, and refs.), and, more recently still, by Mr Broadbent in his book on *Behaviourism* (1961).

Watson's contentions were by no means new. But, like most revolutionaries, he recognized that, without the challenge of an all-out attack, and the emphasis imparted by somewhat sweeping overstatements, his criticisms were likely to attract but scant attention. Nor can anyone seriously doubt that much of his indictment was at that time fully justified, and that his behaviouristic methods have since produced a prolific crop of results in various fields of psychology. Nevertheless, this of itself is not enough to warrant a wholesale taboo on all introspective methods and all introspective concepts. What then are the grounds that are still put forward for these drastic prohibitions?

(1) Of those mentioned by Broadbent, the first, and the earliest in point of time, was the series of apparently insoluble controversies to which the introspective approach had given rise. 'Instead of an ever-growing system of facts and generalizations,' says Watson, 'what introspection has engendered is merely a progressively increasing chaos of conflicting reports and speculative theories.' The outcome, as

James had already complained, was 'not science; it is not even the hope of a science'. Broadbent cites two such controversies as typical—the protracted debate between German psychologists over the possibility of imageless thought and the wrangle among American psychologists over the James-Lange theory of emotions. To psychologists in this country, impressed as they had been by Galton's work on individual differences, such diversified pronouncements seemed only natural: the error lay simply in the fact that each of the protagonists implicitly assumed that his own experience must be typical of everybody else's.

(2) The argument that one most frequently hears today seeks to go deeper, and is based on a special theory of scientific methodology. Science, it is maintained, deals only with events which are 'public', that is, accessible to *any* observer. Conscious states from their very nature can be accessible (if at all) solely to the person who owns them; they are 'private', and consequently 'lack the certainty all scientific data should possess'. What X calls his 'feelings', his 'sensations', and his 'mental images' can be inspected by no one but X himself. All that other psychologists can observe are X's own overt responses to various stimuli—his gestures, his vocal or sub-vocal utterances, his internal visceral changes—in short, X's 'behaviour'. 'The fallacy', says Broadbent, 'in asking a man to describe the contents of his mind is that his words cannot convey directly how he feels. If I say I experience "mixed feelings", *nobody can prove or disprove* the accuracy of my observation. . . . This means that we must reject all concepts that are not defined by any operation.'

Both these arguments are anything but conclusive. To begin with, no scientist outside the field of psychology, not even Bridgman himself, would today support the thoroughgoing doctrine of operationism propounded thirty-five years ago;\* and it would be much easier to quote or construct good operational definitions for the technical terms employed by introspectionists than it would be for some of the newer concepts introduced by contemporary physicists. But this is a minor issue. What I am most anxious to challenge is the widespread assumption that, before any phenomenon can be accepted as a scientific datum, we must (like the Grand Inquisitor in *The Gondoliers*) feel 'no probable possible shadow of doubt, no manner of doubt whatever' about the legitimacy of its claim: in short, that science must insist on certainty. 'There is', says Broadbent, 'no other way besides the behaviourist way of attaining *certain knowledge* about other men' (my italics). But not even 'the behaviourist way' can achieve 'certain knowledge'. Nor would any branch of empirical science be justified in exacting such a guarantee. Dr Taylor (1958) cites Newtonian astronomy. Admittedly, if we start with Newton's assumptions (e.g. with Euclidean space and the law of inverse squares), his purely mathematical conclusions can be deduced with absolute certainty; but the applicability of the assumptions to the stellar universe and the verification of the conclusions must be decided by personal observation; and no observations are infallible. In A.D. 1054 a Chinese 'sky-watcher' at Peiping recorded that he had seen a new star, of such brilliance that it was visible even in daylight, flare up in a corner of the constellation that Western astronomers have christened

\* Psychologists appear to think that operationism is a creed taken over from the physical sciences: physicists regard it as a creed borrowed from psychology. 'In spite of laboratory vocabulary', says an eminent physicist, 'the operational viewpoint is not really a genuine product of physics; it is the echo, in the precincts of physical science, of that curious and typical American doctrine known as "behaviourism"' (A. O'Raahilly, 1938, chap. XIV, 'Operations and Concepts').



the Bull, at the very spot where we now find that powerful radio-emitter known as the Crab Nebula (De Mailla, 1785). Since we cannot unwind the reel of time, no one today can 'prove or disprove' his observation; yet every astronomer accepts it. All that an empirical science requires, or indeed can hope for, is that what an observer reports shall, when duly weighed in the light of what we can infer about his reliability as an observer or reporter, and of all the other relevant facts, seem *highly probable*.

However, the heart of the argument lies in the distinction between what is 'private' and what is 'public'. The distinction is applied at two different levels which are usually confused. First, it is argued that a person's inner experience has to be described in words which he alone understands, i.e. in a 'private' *language*. To some extent this is true. *I* know what I mean by red; but *you* cannot know what I mean by it, since my colour sensations may differ from yours. Nevertheless, though you cannot know for certain, it remains *highly probable* that, unless there is positive evidence for some rather abnormal variation in the brain and sense-organs of one or other of us, your interpretation of the word will be much the same as mine. Thus, even if the experiences are private, the *language* is public so long as it can be understood. Indeed social intercourse would be utterly impossible if a man's words did not 'convey how he feels'.\*

More frequently, however, the behaviourist's distinction refers not to the language, but to the objects or events which the language describes. I and I alone, it is said, can *observe* my inner experiences. Hence, it is maintained, such experiences cannot be data for a 'scientific psychology'. But surely as thus interpreted the distinction is based on the falsest of false antitheses. In no other branch of science would the mere fact that only one man can observe a given phenomenon be offered as a sound *a priori* reason for rejecting it forthwith. Strictly speaking, *every* first-hand observation is necessarily 'private'. Whether certain observations are treated as 'public' turns not on their specific or intrinsic nature, but merely on the context. X is asked: can he or can he not detect a faint point of light somewhere near the centre of a dark field. He answers 'Yes'. If the field is the field of a 4-inch telescope, and the faint point of light is presumed to be one of Jupiter's smaller satellites, then what is observed is said to be 'public'. If the field is the far end of a psychologist's dark-room, and the experiment is intended to determine the subject's threshold for visual sensation, then what is observed is said to be 'private'. Nevertheless, when astronomers are permitted to accept such observations at their face value, why should psychologists be forbidden to do so?

The behaviourist does not wholly rule out experiments undertaken to determine such things as sensory thresholds; but he insists that the threshold shall be regarded as a threshold for 'producing a reaction' not for 'producing a sensation'. And such experiments are sanctioned only when it is possible to make objective checks—e.g. by 'a catch trial... to make sure that the man will always say "I see nothing" when we show him nothing'. In the case of mental images and of 'feelings' (i.e. pleasure, and pain in the sense of displeasure) we have no such means of 'checking the exact boundaries between the situations which a man describes in one way (e.g. by an affirmative)

\* The view that experiences as such are incommunicable seems to be derived from the early pronouncements of certain members of the so-called Viennese Circle; cf. Schlick (1918) and Carnap (1928). But I doubt whether any philosopher would today seriously support it.

and those which he describes in another (e.g. by a negative)'. Here, says Broadbent, lies the 'fatal weakness' of all such introspective studies.\*

Our American contemporaries make much the same points. 'What a man tells us about his emotions or sensations—whether he himself is a psychologist or not—must be viewed, not as the statement of a scientific observation, but merely as part of his total behaviour. . . . Every therapist knows that a patient's statement "I am not angry" is *not* equivalent to the statement "The patient is not angry"; often indeed such verbal behaviour may be used as a datum for the conclusion "The patient is angry".' Introspective statements about feelings, such as 'I feel happy' or 'X feels a pain', we are told, are 'not statements in psychology' or 'statements about genuine psychological variables'. 'The language of the subject and the language of psychology must be kept strictly separate.' For simplicity the psychologist may be permitted to say a human being 'sees' something; but such expressions can only be tolerated on the strict understanding that 'in the language of psychology "see" is defined according to the principle that "S sees O" means, when presented with the stimulus X, S reacts by saying "I see O"' (Mandler & Kessen, 1959, pp. 35f.; cf. Bergmann, 1950, pp. 485f. and Skinner, 1957). For psychology 'S', the subject, thus becomes nothing more than a stimulus-response machine: you put a stimulus in one of the slots, and out comes a packet of reactions.

Meanwhile, what of the patient himself?

In the midst of the word he was trying to say  
He has softly and suddenly vanished away;

for what appeared to be a psychical Snark turns out to be just a mechanical Boojum. But by this time it is, I think, plain that all these reservations and qualifications are little more than linguistic subterfuges. They save the behaviourist's face; but they misdirect our attention. When a patient says that he 'feels a pain', or that, after an injection of morphia he 'now feels happy and comfortable', or again, when a volunteer, after taking a dose of mescaline, declares that she is witnessing 'a glorious panorama of colour and hears the angels sing', or after 40  $\mu$ g of LSD-25 can 'see half-a-dozen gigantic devils chasing a small boy with pointed tridents', surely the crucial feature about the situation is the fact (or should we say the high probability) that the persons in question really do see, hear, or feel what they describe, not that they exhibit this or that 'verbal behaviour'. But, it is objected, 'no words can convey directly' the intrinsic nature of a man's experiences. True; but words cannot 'convey directly' the intrinsic nature of anything. And if 'different subjects use the same vague words to describe different experiences and describe the same experience by different names', that might be said of the pre-scientific language used by the plain man in almost every field of discussion.

However, when psychologists like Myers or McDougall spoke of introspection, they were not thinking of the incidental remarks made by patients to the therapist, but of the reports of skilled and practised observers, trained to use a special technique and equipped with a special scientific terminology—the so-called *systematische Selbstbe-*

\* As we have seen, Broadbent is careful to insist that he does not for a moment wish to deny or ignore the notion that other people beside ourselves can hear, see, feel, or think: he merely holds that psychology is not the proper place to mention or discuss such inner processes. His guiding principle seems to be that of Machiavelli's ideal philosopher: 'Odi, vedi, e taci, se vuoi vivere saviamente'.

*obachtung* developed by Wundt and his co-workers in the psychological laboratory. The words and phrases which Watson and other behaviourists would have us abjure—sensation, after-image, memory-image, feeling, and the like—were used in a technical and somewhat Pickwickian sense, rigorously defined by explicit interpretations. The objection that subjective terms are 'by their very nature obscure, elusive, and ambiguous' has been grossly magnified. Most students find the words that Watson wishes to ban far more intelligible than his own circumlocutions or the recondite phraseology coined by later behaviourists—'sign-gestalt-expectations', 'demanded type of means-or-goal', and so on—even when these are scrupulously defined. Few readers would want their favourite novelist to translate his 'subjective nomenclature' into a behaviouristic idiom; and all that the introspective psychologist seeks to do is to apply the scientist's techniques to the novelist's material, and refine and standardize the novelist's improvised vocabulary.

(3) But it seems clear from various incidental remarks that the most powerful motive influencing the champions of behaviourism in their selection of concepts has its roots in certain tacit assumptions about what kind of concepts are admissible in other natural sciences. The smear-word which Watson most frequently applies when he wishes to deny or deprecate some particular concept is 'unscientific'; and in deciding what is and what is not 'scientific' his criteria are plainly derived from his early training in the physical and biological sciences as they were understood towards the close of the nineteenth century. Psychologists who talked of consciousness and conscious states were, so he contended, 'professing to discover something immaterial and wholly outside the world of natural phenomena'. That world, it was supposed, was essentially a mechanical system, made up of material particles and inexorably governed by a few simple causal laws. Into such a scheme Procrustes himself could not fit the concepts of the introspectionist—concepts such as 'will' or 'purpose' which claim to produce effects in defiance of causation, or such as 'sensations' and 'images' which produce no causal effects at all and whose own causal origin is wrapped in mystery. Accordingly, like Milton's fallen angel, Watson is firmly resolved

To exclude

All spiritual substance by a corporeal bar.

The theory that ensues has all the appeal of a trim, tidy, easily intelligible system, with everything all neatly buttoned up. But unfortunately relativity and quantum theory have meanwhile undermined the whole structure. Today no reputable scientist accepts its oversimplified assumptions. And the physicist himself is the first to emphasize that the schemes he constructs are neither final nor complete: they are models that symbolize reality, not photographs that reproduce it.\*

(4) There is, however, a small but growing band of behaviourists, who now incline to accept this newer standpoint. They prefer to rest their creed on a fundamental principle of methodology which throughout the twentieth century has played an effective part in the formulation of physical as well as of psychological theories—namely, what psychologists call the principle of parsimony and logicians have nicknamed Occam's Razor. As ordinarily stated, it maintains that the number of basic

\* In support of this pronouncement one could cite passages from almost any contemporary physicist—Heisenberg, Schrödinger, Whittaker, Bohr; but I suppose the most recent and authoritative is that of the past President of the Royal Society (Hinschelwood, 1961; cf. also Burt, 1958).



concepts and assumptions should be reduced to the barest minimum. Russell has extolled it as 'the supreme maxim in scientific philosophizing'. 'Minimizing axioms' has long been a guiding precept with mathematical and geometrical theorists; and the practical appeal of the Newtonian system, particularly as re-shaped by the mathematical physicists of the eighteenth and nineteenth centuries, arose largely from the ingenious way it succeeded in embracing a vast and varied mass of observable phenomena in a single deductive scheme derived from an amazingly small array of fundamental concepts and axiomatic 'laws'.

In the field of psychology the most ambitious achievement attained by the application of this methodological principle has been the deductive system outlined by Clark Hull (1952). Yet he himself has recorded how unexpectedly difficult he found it to keep his final list of concepts and postulates down to a plausible minimum. One of his later formulations contains at least eighteen postulates and as many as eighty-six definitions, while the number of theorems deduced from them, and empirically verified, is so far comparatively small. And, since Hull published his last and most comprehensive scheme, the tendency of his behaviouristic successors has been to multiply rather than to reduce the number of their basic concepts.

In my view the attempt to construct a comprehensive and coherent axiomatic system, derived from a minimum list of concepts and postulates, is something that can be reasonably undertaken only when the science in question has, like mathematics or mechanics, arrived at a highly developed stage. Otherwise it almost inevitably leaves out too much. The general trend of modern research, even in sciences which, like astronomy and physics, were once thought to be ripe for a systematic formulation of this kind, has been to reveal an unexpected richness and variety in the universe; and surely the study of human life is likely to be the most complex of all scientific disciplines.

I submit then that none of the behaviourists' arguments really prove what they set out to prove. As a working methodology the cautionary principles advocated by the pioneers of the movement are admirable, particularly in reference to those special fields of research in which they were predominantly interested. But to put them forward dogmatically as an *a priori* limitation to which all psychological theories must in future conform is wholly unjustified and already threatens to block the further advance of psychological inquiry. I am all for a kind of psychological passport officer, who shall scrutinize the credentials of the vast crowd of popular concepts, which, more than in any other science, are apt to intrude from the common discussions of everyday life. But the behaviourist's criteria seem to me as unfair as they are old-fashioned.

However, if anyone proposes, as I have done, to allow the re-entrance of words like 'consciousness', which appear in no other scientific vocabulary, it is incumbent on him first of all to explain precisely what he understands by such terms.

## II. THE PROBLEMS OF CONSCIOUSNESS

*The Nature of Consciousness.* Some years before Watson commenced his onslaught on the concept of consciousness, William James published an essay entitled 'Does Consciousness Exist?' (1902). He reached the surprising conclusion that it does not.

But throughout his discussion he takes it for granted that the word 'consciousness' denotes a special kind of *substance*, or, as he termed it, a 'stuff'. That indeed had been the tacit assumption of most of the older psychologists, including not only those whom James attacked, but James himself in his earlier writings. Occasionally, however, the term was interpreted as the name either of some specific activity or *process*, or of a specific property or *attribute*, of some particular substance, e.g. of the individual organism or of the material brain. And certainly, even though we reject James's own alternative, there can be little doubt that all these earlier interpretations present insuperable difficulties and have proved to be untenable.

Modern scientific arguments are no longer couched solely in terms of substances and their properties or processes; they also include the category of *relations*. And the view I myself have put forward (Burt, 1960 and refs.) is that consciousness in all its forms necessarily implies a specific relation between (a) someone or something who is said to be conscious and (b) something else which he, she, or it is said to be conscious of. When, for instance, I 'see' a rose or 'smell' its scent, when I 'hear' a noise or 'feel' a pain, the verbs that I use really designate various instances of a unique kind of relation for which the most convenient name is perhaps 'awareness'. It is a two-term relation; but, as in so many cases, the nature of the terms is somewhat elusive. Since the relation is asymmetrical, we need a different label for each. Let us provisionally dub them the 'subject' and the 'object', respectively. In the examples cited the 'subject' is indicated by the pronoun 'I'—a word which may denote either what would commonly be called my 'self' or my 'mind', or what the physiologist would call my 'brain', or possibly just the passing thought: its precise character is a matter for psychological examination. The 'object' of the awareness seems in some cases to be a physical 'object', i.e. a 'thing', or at least a part of its surface, as when I see a red tomato, or a silver spoon; in other cases, as when I feel pain from pricking myself with a pin, the mental object seems to be caused *by* the physical object rather than to form part of it. However, my purpose at the moment is not to discuss the problems of perception or the soundness of the sense-datum theory (on that see Burt, 1961*a*), but merely to insist that what confronts us is in all such cases a specific type of relation.

There are only two ways in which the psychologist can avoid introducing some such concept: first, to deny or deliberately ignore its existence, as Watson and the earlier behaviourists decided to do—a procedure so flagrantly inadequate that no competent writer today seems prepared to defend it; or secondly, to reduce all statements containing the word 'consciousness' (or one of its many synonyms) to statements containing some more familiar type of relation. Explicitly or implicitly the notion of 'reduction' is widely adopted in recent psychological writings (cf. Bergmann, 1954, Hempel, 1952, Mandler & Kessen, 1959, especially pp. 114 and refs.); but the notion itself is seldom clearly or satisfactorily defined. By way of definition I suggest the following. If a proposition X is validly reducible to a proposition A, then (i) propositions of class A must be on a lower epistemological plane than those of class X in the sense that they involve fewer concepts, and (ii) X must be logically equivalent to A in the sense that (a) X entails A and (b) A entails X.

In the attempt to reduce propositions about consciousness in this way, a wide variety of alternative relations have been put forward—causation, ownership, resemblance or correspondence, representation or symbolization, the relation of part to

whole, a temporal relation such as 'compresence', and even spatial relations such as 'contact'. In the end, however, each of these reductions seems to issue in much the same type of fallacy. Among psychologists and physiologists the favourite device is a reduction to the relation of causation. A 'subjective statement' such as 'I see a red glow from the fire' implies (so it is said quite rightly) an 'objective statement' such as 'That fire causes a red glow'. The behaviourist usually prefers a rather more elaborate translation, e.g. 'The stimulus, fire, causes me to react with an observable and distinctive type of behaviour—a particular movement or a particular verbal response'. It is then assumed that these causal propositions provide a complete description of the situation, or at least as complete a description as a psychologist is entitled to demand. The crucial test is to reverse the implication. Does the statement 'That fire causes a red glow' (or any of these other versions) really imply 'I see a red glow'? Plainly it does not; some further assertion is needed.

Among contemporary philosophers the commonest mode of reduction is that proposed by Russell and most clearly described by Ayer (1954*a*, pp. 122*f.*). 'To begin with,' he writes, 'we do not accept the realist analysis of our sensations in terms of subject, act, and object. . . . We do not deny that a given sense-content can be legitimately said to be experienced by a particular subject; but this *relation of being experienced* by a particular subject is to be analysed in terms of the relation of sense-contents to one another. . . . Accordingly, we define a sense-content not as the *object*, but as *part of a sense-experience*; and from this it follows that the existence of a sense-content always entails the existence of a sense-experience.' Later he changes the name, and speaks of a relation of *belonging*. When you or I say we 'experience' this or that sense-content, all that we imply (so one gathers) is that this and other sense-contents 'belong to the sense-history of the same *self*'—a phrase which is interpreted to mean 'all such sense-experiences contain organic sense-contents which are elements of the same *body*'.

Now, to begin with, the notion that the 'elements' of which the 'body' is composed are 'sense-contents' is one which only a convinced phenomenalist could accept: it would scarcely be endorsed by the ordinary physiologist. And the change to the notion of 'belonging' is a little puzzling. Since it is said to designate a 'symmetrical relation', it must apparently mean, not 'being a part of which the sense-history constitutes the whole', but 'being a part of some more general whole of which the organic sense-contents are *also* a part'. Let us then apply the same test as before. 'I see a brown dog' may plausibly be said to imply 'There is a brown, dog-shaped content' which is part of a wider whole that includes the further "contents" specified by the theory'; but quite obviously this alternative proposition does not of itself imply 'I see a dog'.

However, most writers who have used the word 'belonging' intend it to designate the relation of possession or 'ownership'. Now in popular parlance we frequently use such expressions as 'I have a pain in my finger', where the verb 'have' could be taken to mean 'I (that is my body) possess a pain'. But a phrase like 'I have a red glow' cannot possibly be used in this fashion. The reason is plain. This idiomatic use of 'have' is only permissible where the following noun itself implies some kind of relation. We can speak of 'having a meal' or 'a wash', because the nouns themselves imply the actions of eating or of washing, and such actions involve relations. Similarly the word 'pain' connotes something that is consciously felt. But with other types of



noun we are obliged to insert a word which explicitly indicates some sort of consciousness: we are therefore forced to say 'I have the *sensation* of a red glow'. Ayer himself in this context always speaks of '*sense contents*'. Thus the equivalence of the two types of relation is really effected by covertly transferring the notion of consciousness from the word which explicitly describes the *relation* in the first statement to the word which implicitly describes the *relatum* in the second.\*

After all, consciousness in the sense of 'being aware of' is essentially a form of knowledge.† Indeed it is the most fundamental form of knowledge: without it observation, which every scientist including the behaviourist automatically postulates, would be impossible. When I am aware of a red glow over there, I know that something is (or at least seems to be) over there, and I also know its distinctive nature, namely, that it is a glowing red. Since knowing is neither a substance nor an attribute, the only logical category under which we can subsume it is that of relation. Hence to deny that there is such a relation between the knower (i.e. me) and the known (i.e. the red glow) is tantamount to denying the very possibility of knowledge. Indeed, it lands us in the self-contradictory situation of saying 'I know that I can never know'.

The arguments that I have used in refuting the attempt to reduce the relation of awareness to a relation of causation, or of part-and-whole, or of belonging, can be used of all the other modes of reduction that have at one time or another been proposed. And, to clinch the matter, we may note that, as Russell (1921, p. 76) rightly points out, 'we are not only aware of things, we are often aware of being aware of them'. Should anyone ask whether I am sure that I see a red glow, I can confidently reply, not only that I am seeing it, but also that I *know* that I am seeing it. I conclude then that the concept of consciousness in the sense of an irreducible relation of awareness is a concept that we can neither exclude nor exchange, if, as psychologists, we are to give an adequate account of human life and behaviour.

Now it is not difficult to show by carrying the analysis a few stages further that immediate awareness is by no means the only species of cognitive relation: knowing in the sense of *savoir* is a different type of relation from knowing in the sense of *connaître*. Older psychologists used also to include such processes as remembering, imagining, judging, and inferring under the same generic heading. Moreover, it may

\* In a later publication Ayer (1954b) recognizes that it is difficult to avoid defining the phrase 'sense content' or 'sense datum' without introducing some transitive verb like 'seeing', 'hearing', etc., which implies the relational notion of awareness. Accordingly he suggests that we should interpret it by saying it denotes something that 'appears'; and by this (so he presently explains) he means 'something that is sensibly present'. But this substitution of an adverb for a verb no more eliminates the reference to a relation of awareness than the substitution of the phrase 'temporally subsequent' would eliminate the reference to a relation of time. We only talk of something 'appearing' when we think of it as actually or hypothetically appearing 'to someone', where the preposition now indicates the relation. The relational view of consciousness has been criticized both by Russell (who formerly accepted it) and more recently by Ryle—but solely as part of the sense-datum theory. Since I do not accept that theory, their objections are largely irrelevant. I have, however, discussed them elsewhere (Burt, 1960, 1961b).

† Some writers have denied that awareness can legitimately be regarded as a form of knowledge; but that is only because they deliberately define the term 'knowledge' so that awareness is excluded. The dictionaries, however, all recognize the sense I have adopted, as well as the more special sense of 'knowing that'. As has often been pointed out, the English verb 'know' does duty for what in other languages—Greek, Latin, German, French, and Welsh—is expressed by two or more different verbs. It was indeed largely to avoid the side-issues that are raised by using the simpler but more ambiguous word 'know' that the earlier introspectionists substituted the somewhat pedantic term 'cognize' or 'cognition' to denote the more comprehensive concept.

be plausibly argued that there are other genera of conscious relation besides the cognitive: there are, for instance, 'affective' relations such as being pleased with, and 'conative' relations such as wanting or desiring. Situations in which one or more of these conscious relations occur differ profoundly from those studied in the physical sciences. Such situations therefore must be studied in and for themselves. When so studied, they display certain regularities, and are in some measure predictable. Hence their investigation is plainly a task for science, and the appropriate science will naturally be what has traditionally been termed psychology.

### III. THE NEED FOR INTROSPECTION

In the investigations I have just mentioned four main groups of problems seem to be involved: first, the nature of the various conscious relations; secondly, the nature of the proximate objects of such relations—i.e. what some writers have termed the 'contents' of consciousness; thirdly, the relation of these objects to physical objects; and finally, the nature of the subject of these relations—whether, for example, we are to think of that subject as a material organism or brain, as an immaterial mind or 'psychic factor' (to use Broad's expression), or as some joint or composite entity. To investigate these various problems would scarcely be possible without employing that particular type of observation known as 'introspection'. Introspection is necessary, not only because it brings new questions to the fore, but also because it alone can supply much of the observational data needed to answer them. However, as many psychologists of a younger generation have entered the field since introspective methods fell into disrepute, it seems desirable to justify their revival by citing a few illustrative instances, and then endeavouring to rebut the commoner objections.

To begin with, most of our detailed knowledge about the psychology of sensation, at least so far as man is concerned, was discovered by experimental studies in which introspective methods were systematically employed. Many of the results—notably the sensory thresholds—could doubtless have been reached, a little deviously, by purely behaviouristic techniques. Yet I question whether anyone would have thought of examining such problems had it not been for the introspective interest in the subject—the interest of the painter, for example, in problems of colour vision, and of the musician in those of sound. And there are many unexpected facts which could never have been established by procedures that barred all forms of introspection—e.g. the curious fact that a mixture of red, green, and blue will produce a pure white, or that the fusion of two stereoscopic pictures gives rise to the extraordinary sensation of solidity or depth, not to mention the so-called phi-phenomenon and the many other phenomena embraced by the term *Gestalt*.

The most striking discoveries that we can lay to the credit of introspective methods are those relating to mental imagery—particularly the wide differences between individuals in the types of imagery to which they are prone. Watson himself was almost completely destitute of concrete imagery, and did most of his thinking by dint of what he calls 'implicit speech'. Quite illogically he inferred that all other people must do their thinking in the same fashion—by 'abridged or incipient movements of the tongue and larynx, often too slight to be detectable even by the most delicate instrumental devices' (Watson, 1931). And when we turn to investigate mental



processes on their higher and more complex levels—notably those of creative and of logical thinking—we find the use of introspective methods still more illuminating.

There are no doubt many areas of research—for instance, those of animal psychology—where such techniques are out of the question, and where introspective terms may prove misleading. Nor would I deny that even in human psychology much may be learnt by treating man as just another animal or even as a mere electrochemical machine. Nevertheless, human beings are neither rats nor robots. And where 'internal' self-observation is available as an adjunct to the 'external' observation of overt behaviour, the marriage of the two is bound to be far more fruitful than the adoption of either to the exclusion of the other.

But science is more than the sheer accumulation of facts and problems. The theoretical psychologist wants to understand behaviour; the practical psychologist to forecast and direct it. As Watson (1919) puts it, scientific psychology has two aims—'to predict human activity with reasonable certainty and to formulate principles whereby human activities can be guided and controlled'. Let us then accept his own criteria, and consider how far the introspective study of conscious processes is likely to assist these aims. Once again, space will permit only a cursory glance at a few of the more typical examples.

When I have to report on a child referred for examination as a potential delinquent or neurotic, I should be quite unable to offer any trustworthy prognosis after merely observing his behaviour and noting his 'verbal responses'. My first step is to gain some insight into his private thoughts and emotions. These I infer and interpret partly on the basis of my own self-observation and partly from the results of comparable observations on other children. The more fully the child himself is able to introspect, the easier it is to secure revealing clues as to his conscious and unconscious motivations. And the value of such an approach is clearly demonstrated by the follow-up results obtained from different clinics. For much the same reason it is almost impossible to direct or control the activities of children or young people with genuine understanding unless we adopt introspective methods and introspective concepts. The study of mental imagery, for instance, turns out to be of great practical as well as theoretical importance—notably in educational and vocational guidance. The most striking instance perhaps is to be found in remedial work on disability in reading (for details see Burt, 1962*a*).

The rule that all introspective terms must be expunged from the vocabulary of 'objective psychology' has played havoc with the interpretation of mental tests and mental factors. In the early days of such research no investigator would have claimed that some new test or factor really measured this or that particular ability, until he had first checked his conclusion by securing introspections from his subjects. Today such identifications are commonly made simply on the basis of some hypothetical preconception—often quite differently by different interpreters. And in almost every field of cognitive psychology—in the study of colourblindness, of sensitivity to noise, of artistic creativity and musical appreciation, of the processes of thought and reasoning, to pick but a few of the most recent examples—this obsessive psychophobia leads to the most far-fetched and misleading periphrases merely to express what are essentially psychical characteristics in purely behaviouristic terms. It is, however, on the motivational side rather than the intellectual that the gulf between the introspective



method and a strict behaviourist approach is most conspicuous. 'Will', 'purpose', 'conation', these are the concepts which Watson singled out for his most ruthless attacks. Instead we must 'substitute mechanical explanations for the meaningless jargon of affective and conative processes', and hold fast to the view that man is 'an assembled organic machine'.

'If you think we are waxworks', said Tweedledum to Alice, 'you ought to pay.' And Watson has to pay a heavy price for his strict adherence to mechanistic principles. It makes nonsense of every form of applied psychology. Educational psychology, vocational psychology, criminology, and psychotherapy—all become impossible if we are to look upon men and women, patients and pupils, as mere automata, devoid alike of reason and feeling. Even in laboratory research on animal behaviour it has become necessary to discard the original restrictions; and it is largely around these motivational problems that the controversies between later behaviourists have revolved. Little by little, however, some of the younger members of the school have ventured to reintroduce a few basic concepts reminiscent of the discredited introspectionist psychology, usually disguised by new names or fresh descriptions. The traditional notion of willing, wanting, or striving—'conation', in short—is rechristened 'drive'; a 'purpose' appears as a 'goal'; and instead of the old pleasure-pain principle—the 'algedonic' law—we have a 'law of effect'. We are forbidden to say the 'satisfaction' we feel at the success which has attended a particular action increases the probability that the action will be repeated; but we are allowed to say that 'the nervous process resulting from a "reward" tends to "reinforce" the nervous process which issued in the reward'.

The object of these heroic circumlocutions is, so we are told, to abolish terms implying 'occult notions, such as perception, pleasure, purpose, or conscious state', and to replace them by interpretations that 'imply nothing more than the mechanical operation of physical causes and effects'. But the futility of the attempt is patent. The hypothetical 'nervous processes' or 'excitations' which the interpretation presupposes are far more 'occult' than the conscious states whose mention is taboo'd. After all, feelings of satisfaction and the like are things that everybody can observe; no one has ever observed the alleged nervous processes that are put in their place, nor has anyone the slightest idea of what they are like, how they can be identified, or indeed whether they really exist.

#### IV. SUMMARY AND CONCLUSIONS

I conclude then that behaviourism, both in its original or 'naïve' form and in its later or 'sophisticated' forms (to use Boring's convenient labels), has proved untenable. As a principle of methodology—particularly in certain specialized fields, such as animal psychology—the behaviourist approach has suggested useful experimental techniques and produced valuable results; but as a basis for a general theory of human experience it is hopelessly inadequate. The need to reintroduce the concept of consciousness seems inescapable. It is quite untrue to declare, as Watson does, that the introspectionist 'never tells us what consciousness is, but merely puts things into it by assumption'. The phenomena of consciousness are not doctrinal assumptions; they are undeniable facts which everyone can verify. In its most conspicuous form—

that of direct awareness—consciousness is a unique relation; it constitutes the basis of all observation, including the observations of the behaviourist himself. And the immediate objects of this awareness—the so-called ‘contents’ of consciousness—are the things we know with the highest degree of certitude.

The behaviourists’ chief objection to descriptions and explanations involving a reference to the processes and contents of consciousness is that, even if we concede their existence, such processes and contents are from their very nature ‘private’ or ‘subjective’. However, as we have seen, the popular antithesis between subjective and objective, between what is ‘private’ and what is ‘public’, turns out, when closely examined, to be at once ambiguous and misleading. What is commonly said to be ‘public’ is known solely ‘by description’ (as the logician puts it), i.e. inferentially; and the ‘private’ experiences of other persons are knowable in precisely the same way. Indeed, as modern physicists like Bohr and Heisenberg are the first to assure us, ‘all the raw material of our knowledge consists of conscious events in the lives of separate observers’.

The behaviourist, as we have seen, goes on to argue that, for purposes of scientific study, what are usually regarded as mental events can always be reduced without remainder to terms of physical events—‘responses of the organism such as can be described in the universal language of natural science’. By ‘responses’, it would appear, the behaviourist originally meant gross bodily movements such as make up what we should ordinarily call behaviour. With this limited interpretation, however, the proposed reduction is plainly impossible. We need to know what intervenes between the stimulus and the subsequent response; and here, so I have argued, introspection yields the most important clues. To avoid introspective concepts, however, later behaviourists preferred to postulate, in addition to ‘overt’ or ‘explicit’ responses (‘macroscopic behaviour’, as I have termed it), certain types of ‘internal’ or ‘implicit’ response—including under this phrase the molecular or submolecular processes of the brain or nervous system (‘microscopic behaviour’). But, even if we could demonstrate that there was in every case a complete one-to-one concomitance between conscious processes and the corresponding nervous processes, we still could not *identify* conscious processes with their neural counterparts. We should still require a set of psychophysical laws stating how, in the various cases, the two were in fact correlated.

In order to study how the various phenomena of consciousness are related to physical processes introspection is essential. Nevertheless, since behaviour and experience are both influenced by unconscious as well as by conscious processes, introspection alone is not enough. The arguments advanced by behaviourists to condemn introspective techniques and terminology have no doubt served a useful purpose by drawing attention to their occasional misuse and their inevitable limitations; but they fail to justify any sweeping prohibition. The proposal of later behaviourists to permit introspective phraseology as part of the subject’s ‘verbal behaviour’ while excluding it from the language of ‘scientific psychology’ misses the vital point of the subject’s report; and such a convention makes it meaningless to ask whether the subject is speaking the truth and impossible to discover what he really intends to assert.

The common complaint that ‘the efforts of the earlier introspectionists produced nothing but controversy and confusion’ is thus a reason not for abandoning intro-

spection, but for greater skill in its use. Rightly understood, self-observation differs in no essential respect from other modes of observation. And even if introspection confronts us with a series of baffling problems, it has also shown that conscious processes exhibit certain regularities, and obey certain laws. It has already revealed a number of suggestive generalizations—some connecting conscious events with physical causes and with physical consequences, others connecting conscious events with one another. The phenomena of consciousness are thus clearly amenable to scientific research; and they are surely of sufficient interest and importance to deserve systematic study in and for themselves. Such a study (as the behaviourist himself reminds us) is undertaken by no other science. It should therefore form the most distinctive feature of the psychologist's task.

I conclude then that observation of self and observation of others are both indispensable. Unless we study our own inner consciousness we cannot fully understand the behaviour of others; and unless we observe others we cannot fully understand our own.

Willst du dich selber erkennen, so sieh wie die andern es treiben;  
Willst du die andern verstehn, blick' in dein eigenes Herz.

### REFERENCES

- AYER, A. J. (1954a). *Language, Truth, and Logic* (10th ed.). London: Gollancz.
- AYER, A. J. (1954b). *Philosophical Essays*. London: Macmillan.
- BERGMANN, G. (1950). Semantics. In V. Fern (Ed.). *A History of Philosophical Systems*. New York: Philosophical Library.
- BERGMANN, G. (1954). Sense and nonsense in operationism. *Sci. Mon. N.Y.* **59**, 140-8.
- BROADBENT, D. E. (1961). *Behaviourism*. London: Eyre and Spottiswoode.
- BURT, C. (1958). Quantum theory and the principle of indeterminacy. *Brit. J. Statist. Psychol.* **11**, 77-93.
- BURT, C. (1960). The concept of mind. *J. Psychol. Res.* **1**, 1-11.
- BURT, C. (1961a). The structure of the mind. *Brit. J. Statist. Psychol.* **14**, 145-70.
- BURT, C. (1961b). The psychology of perception. *Brit. J. Statist. Psychol.* **14**, 173-80.
- BURT, C. (1962a). *Mental and Scholastic Tests* (4th ed.). London: Staples Press.
- BURT, C. (1962b). The sense datum theory. *Brit. J. Statist. Psychol.* **15**, 138-64.
- CARNAP, R. (1928). *Der Logische Aufbau der Welt*. Berlin: Engelmann.
- DE MAILLA, MOYRIA (1785). *Annales de l'Empire Chinois traduits de Tong-Kien-Kang-Mou*. Paris: Alcan.
- HEMPEL, C. G. (1952). Fundamentals of concept formation in empirical science. In *International Encyclopaedia of Unified Science*, Vol. 2. Chicago: University of Chicago Press.
- HINSHELWOOD, SIR CYRIL (1961). *The Vision of Nature*. Cambridge: University Press.
- HULL, C. (1952). *A Behavior System*. New Haven: Yale University Press.
- JAMES, W. (1902). *Essays in Radical Empiricism*. New York: Macmillan.
- MANDLER, G. & KESSEN, W. (1959). *The Language of Psychology*. New York: John Wiley.
- O'RAHILLY, A. (1938). *Electromagnetism: A Discussion of Fundamentals*. London: Longmans, Green and Co.
- RIVERS, W. H. R. (1920). *Instinct and the Unconscious*. Cambridge: University Press.
- RUSSELL, B. A. W. (1921). *The Analysis of Mind*. London: Allen and Unwin.
- SCHLICK, M. (1918). *Allgemeine Erkenntnislehre*. Berlin: Engelmann.
- SKINNER, B. F. (1957). *Verbal Behavior*. New York: Appleton.
- TAYLOR, J. G. (1958). Experimental design. *Brit. J. Psychol.* **49**, 106-16.
- WATSON, J. B. (1913). Psychology as the behaviorist views it. *Psychol. Rev.* **20**, 158-67.
- WATSON, J. B. (1919). *Psychology from the Standpoint of a Behaviorist*. Philadelphia: Lippincott.
- WATSON, J. B. (1931). *Behaviorism* (2nd ed.). London: Kegan Paul, Trench, Trubner and Co.



## AN EXCEPTIONAL TALENT FOR CALCULATIVE THINKING

By IAN M. L. HUNTER

*Psychology Department, Edinburgh University*

This study explores the highly exceptional 'lightning calculation' of a distinguished mathematician who has considerable understanding of his own calculative thinking. Each calculation is a temporally co-ordinated, rapidly flexible onleading which is both unitary and complex. Biographically, it derives from prolonged and intensive practice fostered by circumstances in his upbringing and made possible by a large cognitive capacity which also manifests itself in other forms of intellectual achievement. During calculation, there are 'leaps' of varying compass: there is also notable absence of sensory-type awareness. Ongoing proceeds through apprehending multiple attributes of the presented problem, deciding on some convenient and often ingenious calculative plan, and rhythmically implementing this plan while carrying through opportunistic telescoping and verifying of the ongoing activity.

### I. INTRODUCTION

Prof. Alexander Craig Aitken, F.R.S., was born in New Zealand in 1895. He is a mathematician of recognized distinction who, since 1946, has occupied the Chair of Mathematics in Edinburgh University. He also calculates mentally with a skill which possibly exceeds that of any other person for whom precise authenticated records exist—samples of such records are to be found in Bidder (1856), Scripture (1891), Binet (1894), Mitchell (1907), and Jakobsson (1944). Although Prof. Aitken's calculative skill has gained him a reputation both in Edinburgh and beyond, he rightly values it less than his more high-level, complexly creative intellectual accomplishments. It is a professionally useful side-line which he is reluctant to sensationalize, and only one of his published papers refers extensively to it (Aitken, 1954). This paper contains a talk given to the Society of Engineers on the art of mental calculation. It also contains reports of calculations proposed by the audience during the talk. One of these reports may be quoted to indicate the high order of mental skill under consideration.

'Dr [H. G.] Taylor here asked for the squares of the three-digit numbers 251, 299, 413, 568, 596, 777 and 983, each of which was correctly given almost instantaneously, 568 and 777 taking a little longer. Dr Taylor then proposed the four-digit numbers 3189 and 6371; in each case the square was given in about 5 sec., the lecturer making a momentary error and correcting it in the first case. . . Dr Taylor here proposed [for square root] several of the previous numbers, namely, 251, 299, 413, 596, 777. In each case the square root was given in 2 or 3 sec. to five significant digits, with the remark that for 299 and 596 the last digit might be in excess, which it was. Dr Taylor then proposed [for square root] the four-digit numbers 3189 and 8765. In each case, the result was quickly given to five digits' (p. 298).

The aim of the present study is simply to consider one relatively restricted class of Prof. Aitken's calculative thinking and try to discover as much about it as possible: this class of thinking is decimalizing, that is, transforming a numerical fraction into a decimal. No attempt is made to advance any particular theory of thinking nor to give a rounded psychological portrait of this thinker whose intellectual achievements are remarkable even outwith calculation and mathematics. The aim is naturalistic description of a limited class of highly developed thinking activity. The study was

conducted by interviews and correspondence over a period of a year. Numerical problems were presented to Prof. Aitken and his solving of them was discussed at length. Some interviews were brief, a matter of minutes; others lasted an hour or more; some were informal; some were tape-recorded. It is unfortunate that the tape-recorded interviews occurred at a time when Prof. Aitken was admittedly tired and clearly not in best form. The following extract from a tape-recording illustrates the rapid accuracy of decimalizing even at a time when he is not 'in form'.

*Problem 1.* Decimalize 3 over 408. Pause of 6 sec. during which the data are slowly repeated. 'Point 0, 0, 7, 3, 5, 2, 9, 4, 1, 1, 7, that's as far as I can go.' These digits are spoken at a uniform rate of one every second. He is asked if he has reached the repeating period. 'No, I could go on to it, I think, now. Yes, 1, 1, 7, 6, 4, 7, 0, 5, 8, 8, 2, 3, now that must be it. 3, 5, 2, 9, and so on. The repeating period is sixteen digits.' Again, these digits are produced at a steady rate of one every second. It was later remarked that this was a clumsy problem to present, for it was only  $1/136$ .

*Problem 2.* Decimalize 2 over 63. Pause of 4 sec. during which the problem is repeated. 'Point 0, 3, 1, 7, 4, 6, then the whole thing repeats.' These digits are spoken at a steady pace of two per second.

*Problem 3.* Decimalize 3 over 78. Pause of 4 sec. during which the problem is repeated. 'Point 0, 3, 8, 4, 6, 1, 5, and those last six digits, from the 3 to the 5, repeat.' Those digits are given at a steady rate of two per second. Later, it was remarked that this problem was simply  $1/26$ .

*Problem 4.* Decimalize 1 over 752. Pause of 9 sec. during which the problem is repeated. 'Point 0, 0, 1, 3, 2, 9, 7, 8, 7, 2, 3, 4, that's as far as I can get.' These digits are spoken at a steady pace of about one per second with a speeding up over the last four digits.

*Problem 5.* Decimalize 1 over 57. Pause of 3 sec. during which the data are repeated. 'Point 0, 1, 7, 5, 4, 3, 8, 6, and I can't carry it any further.' He is asked why he does not take it further. 'Well, it's just tiredness just now. If I were fresh, I could carry that further, I think. For instance, I said 6 and I know now that it's 59. You see, I was beginning to approximate. It's 43859, I know that much now.'

These and similar tape-recordings reveal some features of the calculating. Most obviously and impressively, they show the speed of attaining solution and the ease with which an interrupted calculation can be resumed after a lapse of time (when the thinker says he cannot carry an answer further, he does not necessarily mean that he is literally unable to continue but, rather, that he feels it unnecessary to proceed further). These records show, too, the characteristic initial pause and the relatively uniform pace of speaking the answer; also that, according to the nature of the problem, there are variations both in length of pause and in pace of answering. Other recordings strongly indicate that pause and pace are affected by warming-up and fatigue but that, apart from this, pace is a stable characteristic of answering any particular problem, i.e. when the same problem is presented at different times and worked by the same method, the pace of answering is much the same from one occasion to the next. When the thinker reports that a problem was especially difficult to solve, the pace of answering is usually found to have been irregular.

Apart from the above points, these records of calculation in progress reveal little else about the nature of the thinking involved; particularly as the thinker uses no gestures but sits, relaxed and still, while calculating. So it is necessary to ask for introspective reports and explanatory comments. Fortunately, these introspections and elaborations are highly informative. For example, problem 4 was done by taking 752 as 8 times 94, then doing a 94th and dividing the outcome by 8, running the two divisions together. Here is an example of two related commentaries which were submitted in writing.

*Problem 6.* Decimalize  $1/851$ . 'The instant observation was that 851 is 23 times 37. I use this fact as follows.  $1/37$  is 0.027027027027... and so on repeated. This I divide mentally by 23. [23 into 0.027 is 0.001 with remainder 4.] In a flash I can see that 23 into 4027 is 175 with remainder 2, and into 2027 is 88 with remainder 3, and into 3027 is 131 with remainder 14, and even into 14,027 is 609 with remainder 20. And so on like that. Also before I ever start this, I know how far it is necessary to go in this manner before reaching the end of the recurring period: for  $1/37$  recurs at three places,  $1/23$  recurs at twenty-two places, the lowest common multiple of 3 and 22 is 66, whence I know that there is a recurring period of 66 places.'

Six weeks later, he reports on the decimalizing of other fractions which have 851 as denominator. 'Certainly  $15/851$  and  $17/851$  would produce different reactions. For the first, I should have started at once to divide 0.405405405... by 23, three digits at a time. [This is the same basic procedure as before.  $15/37$  is 0.405405 repeating.] The wiser course with  $17/851$  would be to do the corresponding thing upon 0.459459459... [this is  $17/37$ ]; but the fact that 17 into 851 is so very nearly 50 would make me want to correct 0.02 by subtracting  $1/851$  of it, in my own way of rapid compensation. In fact I should have 0.0199765 in a flash; but could not continue so easily, since dividing by 851 and adjusting is not so easy as dividing straight on by 23. I spoke before about not changing horses in mid-stream. Here is a case where the choice would be made at once; I should in fact go ahead with division of 0.459459459... by 23, perhaps deciding to divide two digits at a time instead of three.'

In brief, Prof. Aitken was, for the purposes of this study, posed two kinds of problem. The primary problem was calculative, that is, thinking was directed to attaining a numerical solution. The secondary problem was introspective and explanatory, that is, thinking was directed to attaining a description of primary solving activities. An absolutely complete description was never possible: in part because much calculative activity was unavailable to self-observation, in part because much that was available could not be put readily into words. Nevertheless, these introspections and elaborations were far from meagre and appeared markedly self-consistent. In contrast to many 'lightning calculators', this thinker is a highly accomplished and creative man with wide general interests, he is accustomed to communicating with other people about abstract matters, and is professionally interested in studying mathematical and calculative thinking. So after completing a calculation, he is able to describe its main outlines and some of his subjective experiences; and, if need be, he can supplement this by re-working all or part of the problem, or by working a similar problem. Thus, throughout this study, heavy reliance was placed on what the thinker had to say about his own thinking.

In considering these self-descriptions, some distinction must be made between formal and experiential description. The above report on problem 6 exemplifies a formal report: it spells out a play by play sequence of events in terms of a socially shared language. It is informative, but it is also incomplete and even misleading, for the relation between it and the actual calculation is rather like the relation between verbal instructions for doing a conjuring trick or tying a shoe-lace and the deft execution of these activities. Experiential reports refer to the thinker's experiences during calculation, e.g. to visual imaging if any. They are more difficult for this thinker to achieve; they are also more difficult to interpret. The next section emphasizes formal self-descriptions, and experiential reports will be considered later.

## II. FORMAL CHARACTERISTICS OF MENTAL CALCULATING

Each calculation proceeds, from start to finish, as a temporally extended, ongoing unity: it is a progressive unfolding, opening out, *leading on*. From the moment of



encountering the problem, there is continuous, dynamic extending of the numerical data. It cannot be over-emphasized that this onleading is a unitary activity without distinct break—albeit a unity which is complex, extended in time, and markedly rhythmic. In the interests of analysis, it is possible to consider now this, now that facet. It is especially convenient to isolate three main focal regions; but it must be stressed that these foci are artificially abstracted from a many-sided onleading in which the thinking of each moment is intimately bound to what has preceded and what is predetermined to follow.

In so far as these focal regions can be discussed as separate phases of the dynamic onleading, they typically occur in the following sequence. (1) Apprehending the problem. Here the class of problem (e.g. decimalizing) and the numerical data (e.g.  $1/43$ ) conspire to direct the initial mode of onlead. (2) Deciding the calculative plan. This phase emerges directly from the first and merely carries apprehending further by setting the main course through which leading on will continue. (3) Executing the plan. This leads on further still.

Throughout these three phases, there are some striking general features. There is intimate familiarity with the constraints of calculative language: a rapid ease of converting one set of numerical attributes into other sets; a large repertoire of readily initiated calculative plans along with an appreciation of their implications for the problem in hand. There is large cognitive capacity: ability to carry through a multiplicity of interrelated activities in a short space of time, to act complexly and still hold pertinent data securely ready for use when required. There is the use of rhythm to sustain and co-ordinate this complex onleading. Lastly, there is what might be called sense of propriety: a flexible, opportunistic deploying of resources, an ever-present feeling for what is most fit with regard to the immediate mode of leading on, the future outcomes of super-ordinate plans, the rhythm of thinking; all this in relation both to the data and to his own capabilities.

The three foci mentioned above will now be discussed separately in more detail.

### (1) *First phase: apprehending the data*

A number is apprehended as a multiplicity of numerical attributes and, so to speak, as bristling with signalling properties. This apprehending is immediate, simultaneous, and often autonomous. Regarding immediacy, the attributes issue from the thinker's activities, that is, his perceiving or interpreting of the data; yet there is no awareness of activity any more than most people are aware of activity in, say, seeing two objects as being at different distances from them. Regarding simultaneity, these attributes are apprehended all at once. A person glimpses a short, fat, swarthy, grinning man, and apprehends these attributes simultaneously. A Japanese paper flower is placed in water, and unfolds in several directions, continuing through a variety of co-developments. So it is, for this thinker, with attributes of a presented number.

This simultaneous, immediate apprehending of numerical attributes is often autonomous in that no specific preparation is necessary. For example, on one occasion the thinker heard the year 1961 mentioned, and apprehended this as 37 times 53, and 44 squared plus 5 squared, and 40 squared plus 19 squared. He does not have to set himself to apprehend numbers thus; rather he must set himself to prevent such

apprehending. 'If I go for a walk and if a motor car passes and it has the registration number 731, I cannot but observe that that is 17 times 43. But as far as possible, I shut that off because it interferes with thought about other matters. And after one or two numbers like that have been factorized, I am conditioned against it for the rest of my walk.' He sometimes finds himself squaring numbers which he sees, for example, on the lapels of bus conductors: 'this isn't deliberate, I just can't help it'. Such autonomous activity was certainly absent before the age of fourteen years and exists now as a result of long persistent curiosity about numbers. 'It is a good exercise to ask oneself the question: what can be said about this number? Where does it occur in mathematics, and in what context? What properties has it? For instance, is it a prime of the form  $4n + 1$ , and so expressible as the sum of two squares in one way only? Is it the numerator of a Bernoullian number, or one occurring in some continued fraction? And so on. Sometimes a number has almost no properties at all, like 811, and sometimes a number, like 41, is deeply involved in many theorems that you know.' He has pursued this sort of exercise for half a century, deliberately and enthusiastically in the earlier years and later out of habit and the circumstances of his mathematical occupation.

The readiness with which a presented number leads on to numerical properties is illustrated in the following excerpt from a tape-recording. For each number, the task was to state whether it is prime or, if not, to state its factorization. In each case, there was no clearly perceptible pause between question and answer, and certainly no pause long enough to be measured by stop-watch.

'Q. 963. A. Nine goes into that, so it's too simple to bother with. Q. 386. A. Twice 193: don't have any even ones for they factorize at once. Q. 113. A. Prime. Q. 719. A. Prime. Q. 533. A. 13 times 41. Q. 391. A. 17 times 23. Q. 871. A. 13 times 67. Q. 1211. A. 7 times 173. Q. 313. A. Prime. Q. 417. A. 3 times 139. Q. 529. A. 23 squared: first time [in the entire session of some 30 min.] you've struck a square.'

Although numbers are cues for cognitive action which is often autonomous, it does not follow that apprehending is uninfluenced by context. The way a number leads on varies with the task and with the numbers presented alongside it. Consider the presented number 851 (see problem 6 above). When asked to decimalize  $1/851$ , the immediately dominant attribute is the factorization 23 times 37. When asked to extract the square root of 851, the immediate lead on is that 851 is 29 squared plus 10. When asked to decimalize  $17/851$ , the immediately dominant property is that this fraction is very nearly 0.02. When asked to memorize 8-5-1 as part of a memory span task, there is no numerical leading on: in this situation the separate digits are memorized in a plain, raw sequence.

In short, almost any presented number is rich in meaning, immediately leads on to a constellation of numerical attributes. Some numbers have more attributes than others and the attributes which are apprehended vary with the context—just as, for most people, the same word is apprehended in different ways according to its verbal and situational context.

## (2) *Second phase: deciding the calculative plan*

Anyone who is able, say, to decimalize  $1/23$  plans his calculating in the sense of initiating a schematic and predetermined sequence of activities which, if correctly

executed, lead to the answer. But most adults can readily initiate only one plan, namely, straight long division. By contrast, this thinker can readily initiate any of several plans. His repertoire of calculative plans is large, and the same problem can be solved in different ways—although for any particular problem, some of these plans are more ‘convenient’ than others. Deciding the plan is the key step of the entire performance for it determines the main way in which thinking leads on from the presented data. The following quotation (Aitken, 1954, p. 300) indicates some of the plans which *could* be used with 23 as denominator; it also outlines a modified method of short division which will be referred to later.

‘One can divide by a number like 59, or 79, or 109, or 599, and so on, by *short* division. Take, for example,  $1/59$ , which is nearly  $1/60$ . Set out division thus

$$\begin{array}{r} 6 \overline{) 1 \cdot 0 \ 1 \ 6 \ 9 \ 4 \ 9 \ 1 \ 5 \ 2 \ \dots} \\ 0 \cdot 0 \ 1 \ 6 \ 9 \ 4 \ 9 \ 1 \ 5 \ 2 \ 5 \ \dots \end{array}$$

Here we have the decimal for  $1/59$ , obtained by dividing 1 by 60; as we obtain each digit we merely enter it in the dividend, *one place later*, and continue with the division.

As another example, consider  $5/23$ . Write it as  $15/69$ . Then proceed

$$\begin{array}{r} 7 \overline{) 15 \cdot 2 \ 1 \ 7 \ 3 \ 9 \ 1 \ 3 \ 0 \ \dots} \\ 0 \cdot 2 \ 1 \ 7 \ 3 \ 9 \ 1 \ 3 \ 0 \ 4 \ \dots \end{array}$$

In fact  $5/23 = 0.2173913043478260869565$ , a recurring decimal with a period of 22 digits. One could equally well have written it as  $65/299$ , then carrying out division by 3, two digits at a time, and entering in the dividend *two places* further along.’ [Notice the rationale of this method. It replaces the required divisor, say 199, by a more convenient divisor, i.e. 200. This replacement is then compensated for by continuing to increase the value of the dividend by a two-hundredth of itself. The greater convenience of the substitute divisor is even more evident when it is remembered that dividing by 200 is the same as dividing by 2 with compensatory precautions concerning the decimal place.]

‘There are other possibilities. For example, the mental calculator is, or should be, very familiar with the factorization of numbers; he should know not merely that 23 times 13 is 299, but that 23 times 87 is 2001. For example  $5/23$  is equal to  $435/2001$ ; and if we note that 435 is the same as  $434.99999 \dots$ , we have another method, in which, as we obtain the digits, we *subtract* them from the dividend, so many places later. Thus in the present case

$$\begin{array}{r} 2 \overline{) 434 \ 782 \ 608 \ 695 \ 652 \ \dots} \\ 217 \ 391 \ 304 \ 347 \ \dots \end{array}$$

For example, 217 from 999 gives 782, which we then divide by 2, obtaining 391; this, subtracted from 999, gives 608; and so on.

‘My aim has been to demonstrate, in these various rather simple examples, some part of the repertoire, the armoury of resource upon which a mental calculator may draw, and in regard to the choice of which he must make instantaneous decisions, and keep to them.’

Deciding the calculative plan occurs in that initial pause which follows the presentation of all but the very simplest problems. It can happen that, after onleading has begun by one plan, the thinker is suddenly aware of another and more convenient plan which he must prevent himself from implementing. However, the plan decided on is, usually, the plan which is later described as being most convenient for him to use with the particular problem concerned. It is in this deciding that the thinking is perhaps most subtle and opportunistic: it is this phase of onlead for which a prescription is most difficult to write. With problems as simple as those being considered in this study, there is little if any conscious awareness of comparing the utilities of alternative plans. Often there is a feeling of conflict and uncertainty, a non-detailed awareness of competing modes of onlead; but the resolution of this conflict does not



involve any conscious provisional working out of this or that plan. The decision can usually be justified after the event but little of this justification seems to enter consciously into the making of the decision. 'I have never thought of assigning a method for "ranking" the convenience of one method as against another. Sometimes, when I have embarked on a calculation and a side-flash of a better method occurs, I think I see that it seems better just because it would have taken fewer steps, would perhaps get me twice as many figures by a more rapid process. But apart from this, I think I see at a glance and without going into deep detail that some method will almost certainly be better than some other.'

It is noteworthy that the number of feasible, as opposed to theoretically possible, plans is greatest with problems of intermediate complexity, more specifically, when the divisor is a two-digit number. With a one-digit divisor, plans other than straight division would introduce unnecessary calculative complexity. With many three-digit and almost all four-digit divisors, plans other than straight division would be too complex to carry through readily as a unit: indeed straight division itself may become laborious to the point of being not worth while attempting mentally.

### (3) *Third phase: executing the calculative plan*

Within this schematically predetermined sequence of activities, there is always alertness for alternative ways of executing subsequences, for telescoping and short-cutting chains of calculation, and for apprehending verifying attributes of the on-going thinking. All this is evident in the following introspective report (submitted in writing) on a decimalization done by the method of modified short division mentioned in the preceding section.

*Problem 7.* Decimalize  $1/43$ . 'As you know, I seize at once on a useful property. In this case, that 43 by 93 is 3999, one less than 4000. At once I begin to divide 93 by 4000, entering the answer at the proper place and continuing with the division. Therefore, I have got 0.0232558.... You will notice that, not bothering further with the position of the decimal point, I have divided 93 by 4, getting 2325; I am adding to this one 4000th of itself—strictly speaking one 3999th, but by the time the little increment has been tacked on, it is indeed one 4000th of what then has been set down. Now I said to you that, while this proceeds, I have flashes from the side, small extraneous checks, verifications, and even hints for telescoping or simplifying. One of these occurs almost immediately above. Note the 558. I instantly observe that it is 6 times 93. Excellent check on my accuracy so far. And as I take leave of it, with a 'glance' (not visual), I am dividing by 4000 (that is, by 4 with proper safeguards as to position) and getting 13953488372.... But again I note in the same way the 837, 9 times 93, checking again as I fly along. Also the 372, 4 times 93, telling me that I am almost there because it is *four* times, and 4 is my divisor. And indeed... 8372093 concludes the period. I am back at my first dividend, *with no remainder*, and so everything will now recur.

'But in actual performance, the answer runs with absolute uniformity. The flashes of recognition and reassurance, indicated by asterisks, pass by like flashes of electric bulbs and cause no distraction whatever. 0.023255813953488372093. Well, of course, I could have memorized this decimal. And perhaps it is memory as much as calculation. But the calculation is just as fast as if it were pure memory, and the two intertwine indistinguishably.'

That verification activity is a habitual component of executing the plan is evidenced again by the following report, this time on a simple multiplication.

*Problem 8.* Multiply 123 by 456. 'I see at once that 123 times 45 is 5535 and that 123 times 6 is 738; I hardly have to think. Then 5535 plus 738 gives 56,088. [Note: the 'location' of 5535 is here adjusted so that, in effect, it is 55,350. This thinker does not 'burden the mind with zeros'.] Even at the moment of registering 56,088, I have checked it by dividing by 8, so

7011, and this by 9 gives 779. I recognize 779 as 41 by 19. And 41 by 3 is 123, while 19 by 24 is 456. A check, you see; and it passes by in about 1 sec. [Note: this check breaks the answer down into the originally given numbers by a route which does not merely retrace the method of attaining the answer.]

Perhaps the most outstanding single feature of this thinker is that he leads on so rapidly through such a multiplicity of precisely constrained activities without losing grip on the total calculative ensemble. This large cognitive capacity is clearly evidenced by the above report on problem 7. Quite apart from the running 'side-flashes', this decimalization involves what most people would regard as two distinct, albeit interwoven, chains of calculation. Yet for this thinker, the whole activity leads on as a mobile diversified unity; at most, he experiences one sequence running alongside or underneath another. Thus in so far as distinct calculative sequences are experienced in the modified method of short division, they are experienced as being literally concurrent. However, in more complex calculations there is awareness of alternation. For example, in leading on from the fraction  $1/752$  (problem 4), a 94th is run and an eighth run on the dividend: here there is distinct awareness of two calculative themes. This explicitly experienced alternating is evidenced in the following transcript of a tape-recording.

*Problem 9.* Decimalize  $1/697$ . Pause of 5 sec. 'Point 0, 0, 1, 4, I can't get it any further.' These digits are spoken slowly and irregularly, their speaking, from 'point' to 'four', requiring 12 sec. [The method was straight long-division.] 'You've given me very difficult ones. Those with three digits, I must say, are very hard. I didn't expect them or I might have worked up a wee bit of practice. Having said it, I see a trick for getting it alright. I noticed that 697 factorizes, that's all. You see, you asked me for 697 and that is 17 times 41. Therefore, I'll try again and say: point, 0, 0, 1, 4, 3, 4, 7, 2, that's as far as I can go on that one.' These digits are spoken at a steady rate of one per second. 'I did that by a different method. I mentally worked out a 41st and divided it by 17 at the same time. I did two things. That's severe by the way. A double process, it's very severe. You see you've got to run one decimal and divide it at the same time by something else. You've got to alternate back and forward. I've got to be aware that a 41st is point, 0, 2, 4, 3, 9, and be dividing that by 17 along. So I must be aware of a train and an operation on that train. And the numbers have got to keep coming steadily in. And also, I've not got to lose the clue, I musn't jump one. That's highly complicated.'

Later, in writing, Prof. Aitken commented on this calculating as follows. 'I am really disappointed at my showing here. I should have been able to keep  $1/41$  in mind as 024390243902439... and divided by 17 much further than I did. It would be a long time before the repeating period was reached, because  $1/41$  repeats at five digits,  $1/17$  at sixteen digits, and the L.C.M. of 5 and 16 is 80, so it would be an 80-figure period.' [Note: it is more economical to divide a 41st, repeating at five digits, by 17 than to divide a 17th, repeating at sixteen digits, by 41.]

In this example, the thinker is clearly aware of alternating between two calculative chains. Furthermore, each chain can, after being in abeyance, be picked up correctly without hesitation. This immediate memory capacity is referred to in Aitken (1954, p. 305). 'Prof. Aitken replied that he was able to put aside in storage for a future occasion a result that had already been obtained. He knew that he would be able to bring it out correctly... He thought this ability to put an answer in storage was what distinguished the calculator from what might be called the man in the street. The man in the street forgot the stages in between.' These remarks are substantiated by the following report based on a tape-recording.

*Problem 10.* Decimalize  $4/47$ . Pause of 4 sec. during which 'four over forty-seven' is repeated. 'Point, 0, 8, 5, 1, 0, 6, 3, 8, 2, 9, 7, 8, 7, 2, 3, 4, 0, 4, 2, 5, 5, 3, 1, 9, 1, 4, that's about as far as I can carry it.' These twenty-six digits are spoken at a fairly steady rate of one every  $\frac{3}{4}$  sec. He

is asked whether he has reached the end of the repeating period and, for 52 sec., there is discussion in which he explains that the repeating point might have come after 23 places but, since it had not, it would certainly be at 46 places. Having explained why this should be, he pauses for 5 sec., then begins to complete the decimalization, speaking the digits at a steady rate of one every  $\frac{1}{3}$  sec. '8, 9, 3, 6, 1, 7, 0, 2, 1, 2, 7, 6, 5, 9, 5, 7, 4, 4, 6, 8, now's the repeating point. It starts again at 0, 8, 5; so if that is 46 places, I'm right.'

Subsequent discussion reveals that the main plan was to transform  $4/47$  into  $68/799$ , and then short divide 8 into 68 with the modifications described already. Leading on continued according to this plan until twenty-six digits had been given. At this juncture, it was evident to the thinker that he would have to produce a total of forty-six digits before completing the repeating period; also at this juncture, he noticed that further leading on could be done by an alternative plan. 'And here I stopped, having noticed that 1914 is 3 times 638 which occurred earlier. For a moment I thought I could recover 638297 and so on, and multiply all this by three at the same time as recovering it. But clearly, this was too ambitious and I halted. Later, I returned to my original method based on 47 times 17 equals 799.' Asked if he had any trouble in picking up this interrupted calculation, he replied: 'I know exactly where I was. I always know that. I never have any difficulty in finding the place.'

(4) *Other classes of calculation.* Aitken (1954) outlines some of his calculative plans for obtaining square roots, cube roots, squares, cubes, and factorizations. And it is these plans which seem to constitute the main difference between these calculations and the decimalizing performances already described. The following report (submitted in writing) illustrates the psychological similarity between the activity of leading on to a square root and that of leading on to a decimalization.

*Problem 11.* Extract the square root of 851. 'I at once perceive that 841 is 29 squared. So 29 is a good first approximation. At once I have noted the remainder 10 and halved it (by my rule) and noted mentally that  $5/29$  is 0.172.... At once, I risk 29.172 as an answer which is almost certainly correct to five significant digits. But already I am off on another track, because 29.172 is nearly  $29\frac{1}{2}$ , that is  $700/24$ . And almost before having formulated the procedure in a rational manner, I have divided 851 by  $700/24$ , that is, multiplied it by  $24/700$ . So I feel (rather than see or hear)  $20,424/700$ . But then some experience tells me that 700 times  $29\frac{1}{2}$  is 20,416.66666.... Averaging at speed 20,424 and 20,416.66666.... getting 20,420.33333, dividing by 700, and placing the decimal point in the proper place—and all of this in one continuous follow-through like a good golf stroke—I have 29.17190476.... And I will trust 29.171904 because already, with my grasp on this, I have moved on to further considerations. I have seen that 20,420.33333... differs from either 20,424 or 20,416.66666... by about one part in 5600 or slightly less; I have mentally squared and doubled the result and got roughly 62,500,000 which I choose in round numbers because I pull out from the repertoire the fact that  $1/16$  is 0.0625. Whence I see in a flash that  $1/62,500,000$  of my previous answer is 0.000000466 or so. This very fine correction, subtracted from my previous, gives 29.17190429 say. And I am satisfied to go no further. I count on doing all of these steps in a total time of less than 15 sec. This example, by the way, was not a 'gift'. Some examples are. For example, the square root of 5187 was once proposed. Well, 72 squared is 5184. Therefore it was a "gift" and, in 1 sec., I had 72.02083 which, by my second approximation, I sharpened at once to 72.020833.'

One final brief report may be added, concerning a large multiplication problem presented to Prof. Aitken some years ago by his own children.

*Problem 12.* Multiply 987,654,321 by 123,456,789. 'I saw in a flash that 987,654,321 by 81 equals 80,000,000,001; and so I multiplied 123,456,789 by this, a simple matter, and divided the answer by 81. Answer: 121,932,631,112,635,269. The whole thing could hardly have taken more than half a minute.'

### III. THE DEVELOPMENT OF CALCULATIVE SKILL

Prof. Aitken's calculative skill may well have been genetically predisposed, for there is a family history of arithmetical accomplishment (see Aitken, 1954, p. 307).



In large part, however, it is the outcome of environmental circumstances and especially of a prolonged and intensive practice at mental calculating which he began at the age of thirteen years. He was not a child prodigy. 'Arithmetic in primary school, since I recall hardly anything about it, must simply have bored me. Possibly, I wasn't taught well. Maybe I simply accepted what the teacher said and did it. When I went up to secondary school, with a scholarship from primary school, I was disappointed at my arithmetic mark. I found that it was only 143 out of 200. I lost 57 marks and if I'd had these I'd have been very high up. As it was, I got a scholarship in spite of that.' So he was not an outstanding arithmetician in primary school and, in his first year at secondary school, he did not win the arithmetic prize. His interest in arithmetic first quickened by an incident during an algebra lesson. 'The master chanced to say that you can use this factorization to square a number:  $a^2 - b^2 = (a+b)(a-b)$ . Suppose you had 47—that was his example—he said you could take  $b$  as 3. So  $(a+b)$  is 50 and  $(a-b)$  is 44, which you can multiply together to give 2200. Then the square of  $b$  is 9 and so, boys, he said, 47 squared is 2209. Well, from that moment, that was the light, and I never went back. I went straight home and practised and found that this reacted on every other branch of mathematics. I found such a freedom. I well remember the stage when I was able to square numbers up to 300 and thought—now that is something! But I was to go far beyond that in future years. And so from the age of what might be  $13\frac{1}{2}$  years, up to  $17\frac{1}{2}$  when I left that school, I underwent what can only be described as a mental Yoga. I tried harder and harder things until, in the end, I was so good at arithmetic that the master didn't allow me to do arithmetic.'

This intensive practice in calculation (strictly speaking, in mental algebra) did not, however, restrict interest in other matters and, through this, result in a one-sided, idiot savant type of specialization. Rather, it fostered enjoyment in intellectual accomplishments of all sorts. 'Arithmetic wasn't my only interest by any means. I was interested in literature just as much, Latin, French, English. But this great freedom suddenly encouraged me to think I had a memory and a calculative power capable not only of arithmetic but also capable of, for instance, literary memory. I suddenly moved away. And indeed, did very well at school except for one subject. chemistry, which bored me. The master's way of teaching was the driest dust boys were ever subjected to. Moreover, he once, noticing me to be so inattentive, gave me six of the cane, which didn't improve matters at all. So I just let chemistry go, and so I didn't get the science prize. But I got every other prize that was to be got.' This breadth and enjoyment of intensive intellectual activity persisted beyond his school-days and into the present day. It is evident in his work as a creative mathematician, in his wide cultural interests, in his musical accomplishments, in his memory feats, in his sensitivity to nuances of conscious experience—as mentioned earlier, he is, calculative thinking aside, a rarely accomplished person.

It is noteworthy that this 'sparking off' experience was algebraic in nature and not merely arithmetical: he regards mental algebra as being on a much higher plane than mental arithmetic and incomparably more rewarding. But the effect of this experience was probably aided by two other circumstances. One was that, as a boy, he had no fear of arithmetic; for although it did not interest him then, he had often helped his father (an excellent arithmetician) with book-keeping after school. The

other circumstance was the family tradition concerning his father's elder brother. This uncle 'was a farmer's son and it fell to him to work out volumes of timber and barrels, and so on. He had perhaps found in a book on mensuration Simpson's formula or variants of it and did them in his head. [Prof. Aitken's] father used to say that his uncle was the best arithmetician he had ever known by far and by far, and this legend or tradition was preserved in the family. [Prof. Aitken] supposed that was why when he became thirteen or fourteen he wanted to be something like his Uncle Tom' (Aitken, 1954, p. 307).

It would be wrong to regard this period of 'mental Yoga' as culminating merely in more rapid execution of elementary operations such as simple multiplication and division. This certainly happened: but repetitive drill was incidental to adventurous, persistent, and diversified intellectual exploration of both numerical relations and of his own calculative abilities; so there was also discovery of new and enlarging routines and of progressively higher-order relationships among numbers, and among classes of problems. A calculative routine, say for square-rooting, would be developed; it would be extended to further examples and, thereby, give rise to novel generalizations which, in their turn, would contribute new rules and procedures. New procedures and 'dodges' would then be combined together into yet more novel procedures of greater intricacy and compass for the solving of hitherto unmanageable problems or for the solving of old problems in more ingenious or more abbreviated ways.

A facet of this calculative exploration may be instanced. He developed left-to-right methods of multiplying, adding, and subtracting, as well as dividing. Numbers, if regarded as written sequences of digits, proceed from left to right. And for most calculative purposes, the important digits are those at the left-hand side. So it is 'natural' to treat each number by starting at its beginning. Furthermore, in mental calculation, there is less to remember by left-to-right methods than by the schoolroom methods of right-to-left (the greater memory work of the latter methods is typically managed by externalizing it into a written record). Usually, in left-to-right procedures, the outcome of each successive phase of calculation is progressively incorporated into a single result which approximates more and more closely to the final answer: there is no subsequent need to recall any early outcome once this incorporation has been made; and at a certain stage, the left-hand digits of the answer can be spoken before the entire calculation is complete. However, despite these advantages, left-to-right procedures may involve the thinker in complex co-ordinations. For example, in multiplying, the main product changes with each new partial product incorporated in it; so the attainment of the answer must necessarily copy behind the calculating, and the running memory span may have to be considerable. It is also necessary, as with all mental calculation, to remember the problem and to ensure the correctness of ongoing work because there is no written record of either the problem or the working.

In order to examine some features of left-to-right calculating, the writer asked some seventy students to work several multiplication problems by a method which, though not used by Professor Aitken, has clear affinities to some of his left-to-right plans. For example, the product of 123 and 456 would be attained as follows:  $(100 \times 400) + (100 \times 50) + (100 \times 6) + (20 \times 400) + (20 \times 50) + (20 \times 6) + (3 \times 400) + (3 \times 50) + (3 \times 6)$ . Each successive product is added to the cumulative sum of all earlier products, so that the answer proceeds from 40,000 to 45,000 to 45,600, and so on. The students found mental multiplication easier to do by this plan than by the mental use of the methods they habitually used when calculating with the aid of an ongoing written record. Yet even after an hour of self-paced practice, the students reported that difficulty or total failure could occur at any of many stages in the calculative sequence. They found difficulty in co-ordinating the subjectively distinct activities of: multiplying two numbers; adding two partial products; retrieving an already attained partial product; and recalling the stage now reached in the overall calculative plan. As practice continued, some students reported the development of occasional 'leaps', that is, an enlargement of the onleading correctly accomplished in a subjectively unitary 'step', e.g. proceeding from one partial product to the next without any awareness

of the normally intermediate steps of specifying two numbers, multiplying them, and adding their product to the preceding partial product. However, even this economy raised difficulties for, lacking confidence in the correctness of the 'leap', the student might then undertake its verification and, thereby, disrupt the precarious continuity of the total thinking sequence. The experiences of these intelligent young adults point up something of the adventurousness, persistence, and cognitive capacity of the young Aitken in developing a rapidly effective usage of left-to-right methods.

Although Prof. Aitken cannot recall with certainty the details of his early calculative development, there was clearly that cumulative, hierarchically organized progression which so pervasively characterizes the acquisition of any comprehensive skill, e.g. learning to use a verbal language, to use telegraphic language, to typewrite, to play chess. 'In the process of constantly extending my ability, during the period when I was doing that, I had to think very much of what I was doing. I found that the gains were cumulative and, so to speak, stratified, in the sense that they formed a deposit sinking deeper and deeper into the subconscious and forming a kind of potential upon which, in certain states, I made drafts at astonishing speed.' With increasing speed, economy, and confidence, he learned how to proceed through relatively circumscribed transitions such as those from a particular number to its square, or from two particular numbers to their product (immediated recall?); likewise, how to proceed through less circumscribed transitions of a more general nature such as those from a problem to the most convenient method of solving it (intuitive judgement?). The gross developmental outcome of all this intellectual effort and synthesis was to carry the young Aitken outwith the normal range of mental calculators. Figuratively speaking, he built up an extensive organized library of calculative resources around which he could find his way with ingeniously methodical rapidity: he could, so to speak, skim round the appropriate sequence of files pulling out apposite calculative instructions of varying levels of generality while, at the same time, implementing these instructions as they come to hand and so progressively specifying the co-ordinates of what is required by way of a confidently correct answer—all this before most people have taken in what the problem is.

After the age of about thirty-five, practice in mental calculating was much less intensive than it had been. The result was that the skill deteriorated. Failure to maintain his skill at concert pitch was due, in part, to increasing demands made on his time by other matters and, in part, to lessened interest in mental calculation as such. 'When I came from New Zealand [in 1923] and first used an arithmometer, even of the antiquated types then available, I saw at once how useless it was, how gratuitously useless, to carry out for myself any mental multiplication of large numbers. Almost automatically I cut down my faculty in that direction, though I still kept up squaring and reciprocating and square-rooting, which have a more algebraic basis and a statistical use. But I am convinced that my ability deteriorated after that first encounter' (Aitken, 1954, p. 303). In addition to lack of intensive and regular practice, he also feels that he is 'no longer in prime condition with regard to the sheer physical stamina' necessary for those elaborate calculative feats (such as squaring a ten-digit number) which gave him pleasure in his late teens and early twenties. 'You really have to be in something like perfect athletic form' to do this well. So his present-day skill derives from the more highly developed skill of early adulthood as maintained by relatively infrequent practice. And present-day practice



is, in contrast to the 'mental Yoga' of earlier years, not engaged in primarily for its own sake. It is incidental to his mathematical teaching and research, and especially to the pursuit of those parts of higher algebra where conjectured theorems may be tested by actual numerical examples. 'There, if several examples prove to give the result which one expects from the theorem, there is the incentive to go ahead with the purely theoretical proof; and the quickest way of disproving a conjectured theorem is to try it out numerically and to find that it is incorrect in some trial examples.' So he is not, as he formerly was, interested in mental calculating for the sake of its challenging length or complexity but more for its utility in the subtler tasks of solving theoretical and practical problems involving number.

#### IV. EXPERIENTIAL CHARACTERISTICS OF MENTAL CALCULATING

Prof. Aitken is not able to give such an account of his experiencing during calculation as would enable the present writer to empathize entirely with him: this writer does not fully know what it would feel like to be Prof. Aitken doing a calculation. Reports on the experiential aspects of calculating can be presented under three headings.

(i) *Non-sensory nature of 'numbers'*. He does not calculate by manipulating experienced representations of numbers which have any distinct degree of sensory realism. There is certainly no visual imaging, nor does there seem to be any auditory or kinaesthetic imaging (all three modes of imaging are familiar to him in other, non-calculative contexts). He reports that he could calculate in visual or auditory terms but that this would greatly slow him down. His complexly constrained, intricately timed onleading is, then, unencumbered by vivid imaging and does not, to any reportable degree, involve sensory-type representations of numbers. His onleading is more economical than this; it is further removed from the cumbersome, developmentally primitive level of sensory-motor activity. His calculative coordinations are more akin to those involved in such activities as sprinting, boxing, typewriting, writing, and playing fast ball games. Champion skiers, swimmers and dancers do not perform by consciously representing either limb movements or the many exigencies which influence their activity from moment to moment. Master instrumentalists, singers, and composers do not execute their skills by manipulating sensory-type representations of musical notes. Fluent conversationalists do not emit their sequences of talk by juggling with experienced representations of words. Neither does Prof. Aitken calculate by manipulating 'sensory numbers': he proceeds with a paucity of sensory awareness and his proceeding issues, at the appropriate time, into vocalizing. From moment to moment, he intends in schematic fashion to do this or that. He can describe what he intends to do but, like the fluent talker or typist, he cannot give what an observer would regard as an adequate description of the intending as such.

(ii) *The main calculative stream*. During calculation, he is not an automaton lacking self-awareness. 'I must be relaxed, yet possessed, in order to calculate well. I believe that conscious and subconscious activities are conspiring or in rapid alternation. I seem to move on several different levels. And last of all, when the result is complete, I return to the normal level of ordinary social contact.' In this state of relaxed absorption, he is acutely aware of what might be called the main stream of

calculative activity and also of the often reported side-flashes. However, he cannot describe the nature of this awareness, even though he can describe his doings by translating them into a socially shared language which has words and symbols for numbers, for methods, and for operations on numbers. This socially shared language is, of course, not used in actual calculation; if it were, the whole activity could not proceed with its characteristic rapidity. Social language is only used afterwards as a means of describing to other people what was done. Like all descriptive languages, it is selective and, in some measure, misleading: 'description to others always modifies somewhat the actual sequence of events'.

(3) *Onleading by leaps*. Much of his thinking involves leaping, in the sense of onleading which is accomplished without any involvement of conscious mediation or conscious lapse of time but yet onleading which, if made by less expert thinkers, would involve elaborate and time-consuming successions of try-and-check activities. For example, with this thinker, as with many other people, 12 is the immediated product of 3 and 4: but unlike most people, the transition from '9 times 12,345' to '111,105' is also immediate for this thinker. Consider also his 'simply seeing in one go' the number 1961 as 37 times 53, and 44 squared plus 5 squared, and 40 squared plus 19 squared. Other leaps concern procedural judgements, that is, diagnosing what method is best to use in calculation. These high-level procedural diagnoses derive from a breadth of past experience which is fully comparable to (and possibly in excess of) that which lies behind the so-called position sense of the chess-master, or the swiftly impressionistic diagnoses made by some experienced physicians, or the intuitive snap judgements made by experts in many fields of science, art, and commerce. Yet other leaps occur between attaining an answer and recognizing either that it is correct or that something is wrong. 'I find that if I do not doubt the result of a calculation, it is usually always correct; but if I have any residual doubt, then some correction is usually required.'

One last introspective report is worth mentioning. When he attains the answer to a problem with rapidity, good timing and a feeling of 'all correct', then he cannot easily say whether he calculated this answer or recalled it—especially if he definitely knows that he has done this particular calculation before. In other words, the activity may lack experiential characteristics enabling him to apply to it one or other of two mutually exclusive labels. This report is a reminder of a familiar fact: people classify one onleading sequence of activity as 'thinking' and another as 'recalling' and do so by criteria which are not entirely hard-and-fast: these two broad categories, useful though they are, shade into each other. This being so, it is not surprising that Prof. Aitken should sometimes find it impossible to distinguish between recalling and thinking. What is striking is that, in his case, the vague borderline between recalling and calculating should straddle a complexity of achievement which, for most people, so evidently implies hard thinking.

## V. AN EXAMPLE OF NUMERICAL RECALLING

Prof. Aitken is capable of many remarkable feats of long-term memory for personal experiences, music, and verbal material in English and Latin (e.g. the writer has been able to check the complete accuracy of a list of twenty-five unrelated words which Prof. Aitken recalled from a brief memorizing session in the Edinburgh Psychology

Department of more than a quarter of a century before). With numerical material, one of his memory feats is the rapid recitation of the value of Pi to a thousand places or more. The characteristics of this recitation are noteworthy because of the intensive and rhythmic co-ordination involved and also the changes produced in this co-ordination when recitation proceeds backwards.

Pi is the ratio of the circumference of a circle to its diameter; it is a non-recurring decimal of great mathematical importance. Some years ago Prof. Aitken memorized this decimal to a thousand places. He did it as a kind of intellectual *tour de force* and acknowledges that it 'would have been a reprehensibly useless feat had it not been so easy'. He ranged the digits in rows of fifty, each fifty being divided into ten groups of five. He then found that if he read over the digits in a particular rhythm and tempo, he could fairly easily memorize a group of fifty digits. 'The learning was rather like learning a Bach fugue.' Apart from this rhythm and tempo, there was no 'interpretation' of the digits, no devising of mnemonic relations. 'Mnemonics I have never used and deeply distrust. They merely perturb with alien and irrelevant association a faculty that should be pure and limpid. Our present civilization, not only urban but rural, full of noise and interruption as it is, offers every hindrance to that relaxed meditation upon which the strength of memory thrives best' (Aitken, 1954, p. 301).

As a result of this learning, and periodic 'brushing up', he is able to recite the first thousand digits of Pi and also able to pick up the digits at any given point within the sequence without having to start at the beginning. The writer has tape-recorded this forward recitation, checking it against a typed copy of the digits, and the following report is based on this recording.

Sitting relaxed and still, he speaks the first 500 digits without error or hesitation. He then pauses, almost literally for breath. The total time taken is 150 sec. The rhythm and tempo of speech is obvious; about five digits per second separated by a pause of about  $\frac{1}{2}$  sec. The temporal regularity is almost mechanical; to illustrate, each successive block of fifty digits is spoken in exactly 15 sec.

The second 500 digits are also errorless but there are hesitations and even a few self-corrections. The basic tempo and rhythm is evident but occasional haltings and fumbings give rise to temporal irregularity. Thus, the time for each successive block of fifty digits is 15, 21, 29, 25, 35, 15, 18, 30, 21, and 18 sec., respectively. The total time is 227 sec. This less secure recall is partly the result of fatigue, for he admits that recalling is a strain. Commenting on this recall, Prof. Aitken wrote as follows. 'The first 550 digits were no trouble. Then I think I ran into some trouble for the following reason. Before the days of computing machines there was a kind of competition between calculators (human, I mean) in seeing how far they could calculate Pi. In 1873, Shanks carried this to 707 decimals; but it was not till 1948 that it was discovered that the last 180 of these were wrong. Now, in 1927 I had memorized those 707 digits for an informal demonstration to a students' society, and naturally I was rather chagrined, in 1948, to find that I had memorized something erroneous. When Pi was re-calculated to 1000 and indeed more decimals, I re-memorized it. But I had to suppress my earlier memory of those erroneous digits, 180 of them; and this, I think, sufficiently accounts for a certain tentativeness that seizes me in that region from about the 551st to the 751st digit. But fatigue had doubtless some-



thing to do with it also; and I am inhibited, too, by feeling how useless the feat is—to have written twelve bars of good music would be far more rewarding.'

He is then asked to recite the last fifty digits backwards and does so correctly in a total time of 34 sec. This recitation is still distinctly grouped into blocks of five digits but the whole activity is slower than before and there are some hesitations. In subsequent discussion, the procedural difference between forward and backward recall emerges. In forward recall, no visual imaging occurs. 'Seeing would put me off. If I were forced to visualize [either here or in calculation], I would be much slower. The strong auditory-rhythmic pattern carries almost like bars of music. It's this funny faculty of neither seeing or hearing.' In backward recall, visualizing is employed. 'I bring the numbers along in blocks of five, and manage to see them as a whole, and read them backwards. But that is very much contrary to my usual practice and it goes against the grain.' Visualizing is also used when he is given a string of digits to memorize and is recalling them either backwards or in some order other than the straight forward order (e.g. in the number-square task where a succession of digits are entered horizontally into the cells of an imaginary square and then recalled vertically or diagonally). In the present study, these are the only performances which he finds it necessary to organize in a way which involves distinct visual imaging.

#### VI. POSTSCRIPT

This study has attempted a naturalistic exploration of one class of activities in one exceptional man, and has succeeded in describing some fragmentary aspects of those activities. It now seems fitting to conclude with a comment from the man whose thinking was studied—a comment which pinpoints the main shortcoming of the whole study. 'However much analysis may seem to lay bare the concurrent activities, they are not explained until they are synthesized again and integrated, and then indeed the analysis is seen to be deceptive and defective to that extent. This, of course, is true of all analysis, psychological or physical. "We murder to dissect", said Wordsworth; or phrased otherwise, the whole is ever so much more than the sum of its parts.'

The writer cannot fully express the extent of his gratitude to Prof. Aitken. He acted as subject and answered a host of questions. He commented on successive drafts of the present paper, and permitted its publication despite the personal nature of its contents. His friendly and patient collaboration provided the writer with uniquely first-hand insights into the intricate texture of highly proficient thinking.

#### REFERENCES

- AITKEN, A. C. (1954). The art of mental calculation: with demonstrations. *Trans. Soc. Engrs., Lond.*, **44**, 295–309.
- BIDDER, G. P. (1856). On mental calculation. *Minutes of Proceedings, Institution of Civil Engineers*, **15**, 251–80.
- BINET, A. (1894). *Psychologie des Grand Calculateurs et Joueurs d'Échecs*. Paris: Hachette.
- JAKOBSSON, S. (1944). Report on two prodigy mental arithmeticians. *Acta Medica scand.*, **119**, 180–91.
- MITCHELL, F. D. (1907). Mathematical prodigies. *Amer. J. Psychol.* **18**, 61–143.
- SCRIPTURE, E. W. (1891). Arithmetical prodigies. *Amer. J. Psychol.* **4**, 1–59.

(Manuscript received 11 September 1961)

## A STUDY OF RELIGIOUS BELIEF

By L. B. BROWN

*University of Adelaide*

In an empirical study of religious belief, measures of strength of belief in an Australian student population were compared with beliefs in other areas and with personality and attitudinal variables.

The correlation matrix shows that the religious belief measures belong together and that religious belief has its best relationships with institutionalization and denomination on the one hand and with authoritarianism on the other. The personality measures of anxiety were found to be related to dogmatism in opinionative and factual matters. In this way Thouless's finding that religious belief statements are assented to more strongly than are statements of fact has been confirmed, and extended.

It is concluded that religious belief is a relatively isolated cognitive system requiring strong social support for its maintenance.

### I. INTRODUCTION

Common observation suggests that individuals differ in both the direction and intensity of their religious beliefs, and that these primary beliefs often influence other functions and characteristics. Differences occur not only grossly between believers and non-believers, but among believers, for example, between Methodists and Roman Catholics. It is not clear, however, which are crucial differences, and whether they depend on the performance of characteristic activities, on personality variables, on values or on social support. Alternatively, the differences may depend on an individual's psychological development leading to the establishment of a religious belief-disbelief system, with implications for beliefs about non-religious matters. Obviously, a knowledge of the functional relationships between religious belief and other psychological characteristics is a prerequisite for any adequate psychology of religion.

Accepting that the strength of belief is a central determinant of religious behaviour, Thouless's (1935) questionnaire seems a suitable technique for its assessment. Thouless in his original report showed that religious belief statements are assented to more strongly than are statements of fact, and the present work was begun as a repetition of that study.

The strength of religious belief is therefore defined operationally through the Thouless questionnaire, which implies that religious belief is a cognitive function with varying degrees of subjective certainty, and relationships with other psychological characteristics. Relationships have been hypothesized between the strength of belief and personality variables, general social attitudes, institutionalization and individualism as well as denominational membership. With the Thouless questionnaire it is also possible to study the influence of religious belief on other belief and opinion systems.

## II. THE MEASURES

The variables in the study were defined by measures incorporated in two questionnaires, administered during two testing sessions separated by six weeks. The questionnaires had the following components.

*Intensity of belief*

This was assessed by the use of Thouless's questionnaire, to which eight additional items were added. The items were broken down to areas defined, following Thouless's procedure, in terms of item content. Thouless's classification of the items has been used when making direct comparisons with his material, but because that classification was thought to have deficiencies it was replaced in the later handling of the data by an empirical one. For this purpose the assessments of ten judges were used, and a criterion of at least 80 % agreement in the allocation of each item was obtained. The areas arrived at in this way were beliefs about Christ and (God, other orthodox Christian beliefs, general religious belief (e.g. 'There are such spiritual beings as angels'), opinion (e.g. 'Sex is evil'), fact (e.g. 'Moses was the author of the first five books of the Bible' and 'Tigers are found in parts of China') and a group of miscellaneous items including, for example, 'The spirits of persons who have died can sometimes communicate with the living'.

Each subject was asked to 'show the degree of certainty with which you believe or disbelieve' each of the statements, on a scale from +3, 'a complete certainty that the belief is true', to -3, 'a complete certainty that the belief is false'. This scale was later converted to a score for each item by a simple arithmetical weighting in which -3 was scored 1 and +3 scored 8, and the score for each area was the mean of the weighted scale values for that area.

Thouless had only one sort of uncertainty, namely, 'the statement is as likely to be true as false', but suggested that another could be included. A second kind of uncertainty, 'uncertainty of conviction as a result of lack of information or interest', was therefore allowed. The first of these was scored 4, and the second 5.

*Personality measures*

These assessments were made by the use of the shortened six-item forms of Eysenck's questionnaires for neuroticism (N) and extraversion (E) (Eysenck, 1958), as well as by the Taylor Manifest Anxiety Scale (M.A.S.) (Taylor, 1953) and the items of the M.M.P.I. Lie Scale (L) embedded in the masculinity-femininity items of the same Inventory, although the M-F scale was not itself scored. The measures used were chosen for their general relevance.

*Attitude measures*

A ten-item form of the California F scale (Adorno *et al.* 1950) had been constructed for this population, so that and Lovibond's modification of previous humanitarianism scales were used to assess some general attitudes.



*Institutionalization and individualism*

From a collection of attitude items previously used by Jeeves (1959), two sets of four which had high face validity were included. The institutionalization items entail a high evaluation of the Church; a person strongly agreeing with them all accepts the Church as a primary point of reference in religious matters. Such people may show 'religious conservatism' as Argyle (1958) uses the expression. These items included, 'The Church is necessary to establish and preserve concepts of right and wrong'. Individualism is probably similar to the 'personalism' referred to by Thouless (1959) and the items included, 'A man ought to be guided by what his own experience tells him is right rather than by what any institution such as the Church tells him to do'. They emphasize an individual's own judgments made independently of an external authority, rather than internalized acceptance of a point of view. For the attitude measures and for institutionalization and individualism, the attitude items were in a single questionnaire, with an instruction to show 'how much you agree or disagree with each one'. Strong agreement was then given a weight of 7, strong disagreement a weight of 1, and the scores for each measure obtained by summing weights from the component items.

In addition to these measures, information on age, sex and denominational affiliation was obtained. This last information was derived in a deliberately off-hand way, the question being put orally in the form, 'If you call yourself a Christian, indicate the denomination to which you belong'. This form was used in the hope of counter-acting any habitual tendency to mention a denomination in order to conform.

Table 1. *Classification of, and weightings for, denomination, compared with scale positions adopted by Rokeach*

Denomination	Assigned score	Rokeach scale	Sample size
Roman Catholic	1	0	19
Church of England	3	4	56
Methodist	4	15	40
Other Non-Conformists	5	—	45
including Lutheran	—	10	—
Presbyterian	—	14	—
Baptist	—	17	—
Miscellaneous	6	—	18
Atheist and 'nothing'	9	—	25
Total	—	—	203

To enable denomination to be a variable in the analysis, the denomination given by each subject was assigned a score. In the scoring it was assumed that Roman Catholics stand at one end of a continuum, with those not mentioning a denomination at the other. The scores allocated to the denominations, placed along this continuum, are shown in Table 1 with Rokeach's (1960) results (available after this work had been done) for comparison. To determine his values Rokeach 'applied an "unfolding method" technique similar to that developed by Coombs (1950)' (Rokeach, 1960, p. 297). There are other methods of classification; Argyle (1958), for example, groups denominations as conservative religion, protestantism, sects, and liberalism, but a direct weighting of denominations was considered to be more suitable for this study.

III. THE SUBJECTS

In work of this kind there are problems of definition, not only of what 'religious behaviour' is, but also of who is 'religious' and consequently who should be studied. At one extreme would be 'religiosity', studied by selecting subjects from 'Pentecostal' and 'Disciples of Christ' sects as did Broen (1957), and at the other there would be no special definition, and so the sample would not be selected for religious bias. The latter course seems to be appropriate when the nature of Christian belief as a cognitive system is being investigated. It was therefore thought defensible to use students for this inquiry. They constitute a reasonably homogeneous group, and in this sample they all shared a similar Judaeo-Christian background, with only a minority (four people) drawn from unorthodox groups, such as Seventh Day Adventists. Table 1 shows the numbers in each denominational group for the main sample.

The Thouless questionnaire was given in 1959, and repeated in 1960, together with the personality and attitude questionnaires. This is the main sample, in which there were 203 first-year psychology students in the University of Adelaide, with a mean age of 22 years.

IV. RESULTS

The data were analysed first for comparison with previous work, particularly Thouless's findings. A factor analysis was then carried out.

*Reliability of the measures*

Reliability of the Thouless questionnaire and of the institutionalization and individualism measures was assessed by a retest after 8 months of forty subjects from the original group, who went on to the second-year psychology course. Although this method of assessing reliability has disadvantages, it is the only one appropriate in this situation. The test-retest correlations were:

Orthodox beliefs	.85	Miscellaneous	.50
General religious beliefs	.92	Institutionalization	.53
Opinion	.35	Individualism	.60
Fact	.30		

The low reliability coefficients for the Opinion and Fact scores are probably spurious, although they are not without interest in themselves; they may be due in part to the instability of beliefs about these questions, and to homogeneity in the scores. There are no significant differences between men and women.

The correlation between the Extraversion and Neuroticism scores is  $-.09$ , which conforms to the correlation of  $-.05$  for these questionnaires reported by Eysenck (1958). These measures were also found to be independent for the males and females separately. The correlation between neuroticism and Manifest Anxiety is  $+.511$ . The attitude measures of authoritarianism and humanitarianism were independent of the personality measures.

Jeeves (1959) found his institutionalized type correlating with anti-semitism more than the other types, and a similar result was obtained here with a correlation of

+·466 between institutionalization and authoritarianism, and +·022 between individualism and authoritarianism. Humanitarianism has low correlations with both of these variables ( $-·169$  and  $+·095$  respectively).

### *Intensity of belief*

Thouless in his first table shows the 'mean certainty of different test statements' and concludes that 'the tendency to certainty is less strong amongst non-religious (statements) than amongst those of a religious order'. Identical results were obtained here, with samples taken in each of two succeeding years. In the main sample the mean certainty for religious statements was 2·10 (Thouless 2·13) and for the non-religious statements it was 1·56 (Thouless 1·575). The detailed results are shown in Table 2.

When the degree of certainty for each item is considered, the results are again essentially the same as those obtained by Thouless. The correlation of mean certainty for each item between Thouless's order and the present one is  $+·65$  (S.E. =  $·16$ ). The small differences that occur can be explained from the item content in terms of a shift of interest after thirty years and a different culture. The differences are not in the religious belief items, but in items like 'The total national debt of Great Britain is more than a thousand million pounds' (40), which showed the greatest difference, 'Members of the leisured class are supported by the surplus value created by the workers' (37) and 'Tariffs improve trade' (38), although none of these differences reaches the  $·05$  level of significance.

It is of some interest that the essential similarity of results occurs despite the addition of a second kind of zero category. Thouless commented that a difficulty with his results was the ambiguity of the zero response, and as has been already stated two kinds of uncertainty were allowed in this study. A further point of interest because of the close similarity are the responses to the item, 'Green is a primary colour' for which Thouless was unable to 'explain the high certainty'. He found a mean certainty of 2·33 for this item, and in the present study the mean certainty was 2·32.

### *Denominational differences in belief*

Thouless used an internal criterion to distinguish between believers and non-believers, but the results set out in Table 2 with the percentages in each of the denominational groups showing 'complete certainty' about the items, make the nature of differences between believers and non-believers clearer. As the primary interest is in the extremes of certainty it seems preferable to examine the proportions who strongly assent to each item rather than to use an over-all mean of certainty.

Looking first at the Group D items, Thouless's 'affectively indifferent non-religious beliefs', it is seen that there are no significant differences between the denominational groups. On the other hand, when the Methodists and Roman Catholics as the most homogeneous religious groups in the sample are compared, the following items in Group A (the religious beliefs) show differences with significance better than the  $·05$  level: 'There is a Hell in which the wicked will be everlastingly punished' (23),



'There are such spiritual beings as angels' (9), 'There is a personal Devil' (5), 'Attendance at church is a better way of spending Sunday than taking a walk in the country' (18) and 'There are spiritual realities of some kind' (3). When Methodists and atheists ('nothing') are compared, it is only in the items, 3, 9 and 18 that they differ significantly from each other. However, the mean percentages with complete

Table 2. *Showing mean certainty of different test statements (after Thouless), and percentages in the denominational groups having 'a complete certainty that the belief is true', or for statements followed by -3, false*

Group*	Item	Mean certainty		Complete certainty				
		Thouless	Brown	Nothing	Others	Metho- dist	C. of E.	R.C.
A	1 God	2.25	2.44	26	76	84	45	83
	2 Christ	2.27	2.43	15	76	87	52	96
	3 Spirit real	2.40	2.22	30	59	55	50	79
	4 Created world	2.19	2.33	11	76	76	45	92
	5 Devil	2.19	1.98	11	47	24	17	50
	7 God power	2.32	2.44	21	76	84	52	96
	8 God good	2.38	2.42	23	79	84	52	96
	9 Angels	1.67	1.86	2	38	20	10	79
	10 Jonah	2.01	1.74	2	32	13	10	25
	14 Spirit after death	2.07	2.24	13	59	67	40	88
	18 Sunday Church	1.82	1.74	0	24	33	2	58
	20 Christianity > Budd- hism	1.80	1.70	6	50	44	19	63
	21 Bible literal	2.44	2.25	0	24	2	2	17
	22 Man responsible	2.62	2.63	64	65	73	62	83
B	23 Hell	2.35	2.10	6	47	16	10	92
	6 Matter (-3)	2.25	2.06	23	44	62	50	83
C	16 God (-3)	2.34	2.55	34	79	95	74	92
	11 Evolved	2.23	2.11	30	29	44	52	8
D	12 God impersonal (-3)	1.89	1.91	21	56	60	29	58
	13 Evil reality	2.32	2.21	19	65	62	50	75
	19 Moses (-3)	1.30	1.24	17	12	31	19	13
	24 Spirits commune (-3)	1.66	1.61	15	38	33	13	25
	26 Evolution	2.17	2.11	26	41	62	50	50
	28 Mary, Queen of Scots	1.22	0.67	6	3	7	10	4
E	30 Tigers	1.23	1.20	6	3	16	2	4
	31 Hornets (-3)	1.21	1.60	19	26	35	33	21
	33 Light speed (-3)	2.01	1.94	47	29	38	33	29
	35 Green	2.33	2.32	21	35	27	31	29
	40 National debt	1.82	0.51	6	6	5	5	4
	17 Expand universe (-3)	0.95	1.14	11	12	15	19	17
	34 Bacon (-3)	1.83	1.84	26	32	42	40	25
	37 Leisured class (-3)	1.99	1.23	2	9	11	5	21
F	38 Tariffs (-3)	1.72	0.98	4	3	4	7	4
	39 India	1.78	1.34	12	12	20	7	4
G	15 Religion opium (-3)	2.07	1.81	13	32	49	12	54
	25 Right triumphs	2.01	1.80	13	50	60	24	79
G	27 Hardship	1.77	1.55	6	18	18	14	0
	29 Relative	1.87	1.39	21	6	13	10	25
	32 Sex evil (-3)	2.48	2.68	74	82	87	71	88
	36 Sunlight	2.83	2.59	68	76	76	67	54

\* The grouping is that made by Thouless (1935). He called the groups: A and B, Religious beliefs; C, Unorthodox religious beliefs; D, Affectively indifferent non-religious beliefs; E, Political beliefs; F, Religious 'tabloids'; G, Non-religious 'tabloids'.

certainty on items in Group A are: 'nothing' 15.3, Methodists 50.8 and Roman Catholics 73.1. The differences between these means are significant beyond the .05 level.

It is clear from these results that although Thouless's findings about differences between religious and non-religious statements are substantiated, when the results are seen in greater detail, a more complex pattern of responses emerges. If the beliefs belong to a formally prescribed religious or theological system they are strongly adhered to by people who accept other parts of the system, and regard themselves as belonging to the institution that sanctions the beliefs. The differences between Methodists and Roman Catholics that have already been noted are reflexions of the eschatological theology of these two churches.

Table 3 shows the responses for 'complete certainty' and 'strong conviction' for the item, 'There is a God who is all-powerful' (7), which is central to any religious position, and also for the item 'Mary, Queen of Scots, was beheaded between 1580 and 1590' (28). The pattern of answers for the first of these is a common one for the religious items, and differs from the pattern in other types of question. Members of the Church of England (see Table 3) do not maintain such a strong degree of certainty about religious items, more of them than of other groups tending to accept 'strong conviction' rather than 'complete certainty'.

Table 3. *Showing the two highest degrees of conviction for item 7, 'There is a God who is all-powerful', and for item 28 'Mary, Queen of Scots, was beheaded between 1580 and 1590' by percentages*

	'Nothing'	Others	Methodists	C. of E.	R.C.	Mean
Item 7						
+3	21	76	84	52	96	65.8
+2	13	12	9	24	4	12.4
Sum	34	88	93	76	100	78.2
Item 28						
+3	6	3	7	10	4	6
+2	9	9	5	14	13	10
Sum	15	12	12	24	17	16

In questions of fact which Christians might be expected to know about, such as 'Moses was the author of the first five books of the Bible' (19), there are few significant differences between non-Christians and Christians. For example, 31 % of the Methodists were certain that this statement was false, while 15 % of all other denominations and 17 % of non-Christians gave this reply. There is a similar pattern in the item, 'Man has evolved from lower forms of life' (11), except that the miscellaneous Christian group and the Roman Catholics are closer together with 32 % and 21 % respectively showing complete certainty compared with only 5 % of all other groups.

#### *Other denominational comparisons*

The assertion that denominational groups differ from one another in personality characteristics and in attitudinal variables has been commonly made, some of the material being summarized by Argyle (1958) in his chapters 8 and 9. The results from the present study are summarized in Table 4, the only significant findings being a higher anxiety score (M.A.S.) and lower authoritarianism (F) score for those not belonging to any denomination. Argyle's summary statement that 'for people between

the ages of 16 and 30 the religious individuals are somewhat more neurotic' (Argyle, 1958, p. 106) is not supported. The finding relating to authoritarianism is not a clear one because of the significant difference between the non-denominational males (mean 27.7) and females (mean 40.3) that is concealed here, and the fact that the mean for these females is close to the mean of the denominational groups. Rokeach reports an equivalent mean on ten authoritarianism items of 37.8 for Catholics and 31.6 for non-believers (1960, p. 110).

Table 4. *Showing mean scores of the denominational groups on the personality and attitudinal variables*

N		Manifest Anxiety	Lie Score	Neuroticism	Extra-version	Authoritarianism	Humanitarianism
25	Nothing	18.0	2.7	8.45	7.0	34	58.3
63	Others	15.6	3.7	7.0	6.3	42.5	55.1
40	Methodists	15.5	3.1	7.1	8.2	40.2	56.2
56	C. of E.	14.1	3.2	7.7	7.0	45.1	53.7
19	R.C.	15.9	3.5	8.1	6.8	43.6	47.7
	Over-all Mean	15.40	3.21	7.45	6.83	42.00	53.66
	S.D.	6.87	2.02	3.75	3.23	11.04	10.21

Table 5. *Matrix of correlations (being output of SILLIAC programme K 2), corrected to two places and with the decimal points omitted*

	M.A.S.	L	N	E	F	H	Christ	God	O.B.	G.B.	Opin.	Fact.	Misc.	Instit.	Indiv.	Denom.	
M.A.S.	-06																
L	-04	-23															
N	-18	+51	-22														
E	-13	-12	+04	-09													
F	-12	+02	+19	+04	+06												
H	+13	+16	+04	-01	+02	-18											
Christ	-34	-01	+22	+02	+12	+43	-19										
God	-22	+02	+15	-01	-02	+41	-10	+76									
O.B.	-28	+01	+15	-05	+03	+44	-14	+82	+86								
G.B.	-22	-01	+16	-07	+01	+41	-11	+73	+68	+76							
Opin.	-15	+18	-03	+07	+04	+24	-02	-04	-01	-02	+01						
Fact.	+01	+05	-03	+08	+03	+08	-03	-07	-11	-11	-01	+30					
Misc.	+06	+10	-05	+06	+04	-05	+04	-09	-17	-11	-07	+09	+23				
Instit.	-13	-12	+14	-04	+08	+47	-17	+47	+49	+45	+11	-01	-09				
Indiv.	+02	+17	-11	+13	+01	+02	+10	-16	-20	-19	-26	+11	-07	+18	-14		
Denom.	+10	+14	-02	-00	-02	-34	+27	-47	-37	-42	-38	+03	+02	+05	-36	+26	
Sex	-35	+03	+17	+09	+14	+21	-16	+27	+26	+24	+13	+06	+04	-07	+19	+02	-15

O.B., Orthodox Belief; G.B., General Belief; Opin., Opinion strength; Fact., Factual certainty; Misc., Miscellaneous items; Instit., Institutionalization; Indiv., Individualism; Denom., Denomination.

### *The factor analysis*

All of the component scores were inter-correlated (using the SILLIAC programme K 2); the resulting matrix is given in Table 5. A factor analysis was then carried out with programme M7, and six iterations were made to obtain stable communalities (using programme K7). Because the factors obtained from this refactored material were readily interpreted, there seemed no advantage to be gained by a quartimax



rotation. (Sanai (1950) adopts a similar procedure.) Thus the factor analysis presented is a principal axis solution with the factors having latent roots greater than 1.00 after six iterations interpreted directly.

The latent root of the first factor is 4.63, and on re-factoring after six iterations the latent root is 4.30. The principal loadings are:

Orthodox Christian belief	+ .894	Individualism	— .278
General religious belief	+ .812	Age	— .365
Institutionalization	+ .656	Denomination	— .581
Authoritarianism	+ .598		

The loadings of the personality variables on this factor are:

Manifest Anxiety	+ .072
Neuroticism	+ .033
Extraversion	+ .074

The loadings of the factual and opinionative scores are:

Opinion strength	— .032
Certainty of facts	+ .061
Miscellaneous items	+ .144

This factor accounts for 43 % of the total variance.

The first is clearly a religious belief factor, which shows religious belief to be a system isolated from the opinionative and factual systems, without a relationship with the personality measures and yet strongly associated with both an orientation to the Church (institutionalization) and to denominational membership. The loading on authoritarianism confirms previous findings, and as might be expected individualism has a negative loading.

The second factor has a latent root of 1.93 and on refactoring the latent root is 1.41. The principal loadings are:

Manifest Anxiety	+ .725	Miscellaneous items	+ .285
Neuroticism	+ .704	Individualism	+ .376
Opinion strength	+ .500	Lie Score	— .312
Factual certainty	+ .339	Age	— .337

This factor accounts for 15% of the variance and is obviously a personality factor of neuroticism or anxiety. Extraversion has a loading of —.019 on this factor, orthodox religious belief a loading of —.009 and general religious belief one of —.015. Authoritarianism has a loading of +.216.

The second factor is similar to one identified by Rokeach (1960, chapter 19), who found dogmatism and anxiety to be part of a single psychological factor, in which authoritarianism is not included. However, Rokeach goes on to report differences in dogmatism and anxiety between groups of Catholics, Protestants and non-believers.

It is of considerable importance theoretically that the opinion and factual scores, rather than the religious belief measures, have high loadings on this personality factor. This is a finding that is contrary to expectations from other studies, for example Funk (1956) found a correlation of +.29 between Manifest Anxiety and orthodoxy of belief in students.

In summary it appears that the strength of religious belief is associated with acceptance and membership of a church, while certainty about opinionative and factual matters is associated with personality variables, and specifically with measures of anxiety.

The remaining factors all have latent roots less than unity after refactoring, and their content in no case adds to the interpretation. The third factor has a latent root of 1.47 on the first factorization and 0.95 after six iterations. It is a bi-polar factor, again showing the relationship between opinion strength and personality, while the fourth factor with latent roots of 1.28 and 0.67 after iteration is an age factor on which factual certainty has a loading of +.405 and orthodox religious belief a loading of +.053. Copies of the tables of unrotated principal components and of residuals after the removal of factors may be obtained from the author.

## V. DISCUSSION

The results of the study point towards religious belief being a relatively isolated cognitive system, in which intensity of belief is independent of the strength of opinions about other matters. The relationships between belief and church membership, attitudinal acceptance of the Church ('institutionalization') and authoritarianism suggest that strong social support is required for the maintenance of a system of religious belief.

Denominational adherence also varies systematically with the strength of religious belief. Historically the Church has maintained its doctrines in ways that differ from one denomination to another, and each church has usually handled heresies vigorously. This study shows that such a system of external controls is paralleled in the internal structure of individuals. There is clearly no outside agency to enforce or test positive religious beliefs apart from some form of the Church, and with this analysis it is difficult to see how a highly 'individualized' person can have strong conventional religious beliefs, which are necessarily institutionalized. In another study by the present writer an open-ended question about attitude to the Church revealed lack of social support as a reason for developing an unfavourable attitude to the Church. This social basis of religious beliefs is confirmed by the absence of correlations with personality factors and the small but negative correlations ( $-.34$  to  $-.21$ ) between belief and age in this sample, which is however both homogeneous and young. It is therefore church membership and attitudinal acceptance of the Church, rather than personality variables, that are related to the strength of belief here. The relationship in older groups remains to be established.

The acquisition of religious beliefs may be explained by the operation of social learning and it might be possible to identify some critical social experiences among believers, as Witkin suggests in another context (quoted by Murphy, 1958). As religious beliefs are learned under conditions of inter-denominational rivalry with varying authorities quoted to support beliefs, and differing uses made of biblical authority, institutionalized religion frequently is inconsistent with liberal attitudes. It is a well documented fact that more religious people tend to be less tolerant (Prothro & Jensen, 1950; Wilson, 1960) and the present analysis shows the same relationship. Argyle's (1958, p. 84) suggestion that it is not the genuinely devout,

but the conventionally religious who are prejudiced is consistent with the present interpretation.

Religious certainty is unrelated to certainty about factual and opinionative matters in this study, while the relationship between anxiety and certainty about matters of opinion is similar to that reported by Rokeach, and also to Brengelmann's (1960) finding from very different material. Rokeach reports scores on dogmatism and anxiety correlating from .36 to .44 in various groups (1960, p. 364), and although the correlation between Manifest Anxiety and opinion strength in this study is +.18, the factorial relationship is clearer than the correlation suggests. There is thus confirmation of a relationship between dogmatism and anxiety.

Ordinarily it would be expected that statements which cannot in principle be tested directly, such as the existence of 'such spiritual beings as angels' might be held less strongly than are statements which can be settled by reference to commonsense experience. However, the results show that these religious statements are held more strongly than the factual ones, perhaps to avoid the uncertainty which would necessarily follow from a weaker acceptance. The results also show that it is easier to be uncertain about a factual matter which can be settled, than to be uncertain about something which is literally a matter of belief; certainty about religious matters is possible because of the social support that can be evoked to sustain these beliefs. Anxiety plays a role in holding matters of opinion strongly, but not in matters of belief.

A cognitive theory of religious behaviour is suggested by this study, but as Argyle shows, there are other theories. The present cognitive interpretation can be extended to include affective components by drawing attention to the attachments to and arousal value of signs, symbols, objects and verbal formulations in any religious system. The function of these things is to maintain appropriate responses and so, for example, it is common for preachers to repeat well-known formulations, and particularly texts, for their arousal value, and not for their cognitive significance.

The author wishes to thank Dr R. H. Thouless for his comments on an earlier draft of the paper. He is also grateful to Professor M. A. Jeeves and Mr S. H. Lovibond for allowing the use of some of their measures.

#### REFERENCES

- ADORNO, T. W. *et al.* (1950). *The Authoritarian Personality*. New York: Harper.
- ARGYLE, M. (1958). *Religious Behaviour*. London: Routledge and Kegan Paul.
- BRENGELMANN, J. C. (1960). Learning and personality. IV. Certainty and output motivation. *Acta Psychol.* **17**, 326-56.
- BROEN, W. E. (1957). A factor analytic study of religious attitudes. *J. Abnorm. Soc. Psychol.* **54**, 176-9.
- COOMBS, C. H. (1950). Psychological scaling without a unit of measurement. *Psychol. Rev.* **57**, 145-8.
- EYSENCK, H. J. (1958). A short questionnaire for the measurement of two dimensions of personality. *J. Appl. Psychol.* **42**, 14-17.
- FUNK, R. A. (1956). Religious attitudes and manifest anxiety in a college population (Abstract). *Amer. Psychol.* **11**, 375.
- JEEVES, M. A. (1959). Contribution on prejudice and religion in Symposium on problems of religious psychology. *Proceedings of the 15th International Congress of Psychology, Brussels*, pp. 508-9. Amsterdam: North Holland Publishing Co.



- MURPHY, G. (1958). *Human Potentialities*. New York: Basic Books.
- PROTHRO, E. T. & JENSEN, J. A. (1950). Interrelations of religious and ethnic attitudes in selected southern populations. *J. Soc. Psychol.* **32**, 45-9.
- ROKEACH, M. (1960). *The Open and Closed Mind*. New York: Basic Books.
- SANAI, M. (1950). A factorial study of social attitudes. *J. Soc. Psychol.* **31**, 167-82.
- TAYLOR, J. A. (1953). A personality scale of manifest anxiety. *J. Abnorm. Soc. Psychol.* **48**, 285-90.
- THOULESS, R. H. (1935). The tendency to certainty in religious belief. *Brit. J. Psychol.* **26**, 16-31.
- THOULESS, R. H. (1959). Discussion of Symposium: Problems of religious psychology. *Proceedings of the 15th International Congress of Psychology, Brussels, 1957*, p. 514. Amsterdam: North Holland Publishing Co.
- WILSON, W. C. (1960). Extrinsic religious values and prejudice. *J. Abnorm. Soc. Psychol.* **60**, 286-8.

### APPENDIX

The items of the attitude measures, measures of institutionalization and individualism and the enlarged Thouless questionnaire grouped together in areas. The numbers in brackets refer to the order of items in the Thouless questionnaire set out in Table 2.

#### *Authoritarianism*

1. Obedience and respect for authority are the most important virtues children should learn.
2. No weakness or difficulty can hold us back if we have enough will power.
3. Every person should have complete faith in some supernatural power whose decisions he obeys without question.
4. What the youth needs most is strict discipline, rugged determination, and the will to work and fight for family and country.
5. Young people sometimes get rebellious ideas, but as they grow up they ought to get over them and settle down.
6. Sex crimes, such as rape and attacks on children, deserve more than mere imprisonment; such criminals ought to be publicly whipped, or worse.
7. People can be divided into two distinct classes: the weak and the strong.
8. There is hardly anything lower than a person who does not feel a great love, gratitude, and respect for his parents.
9. If people would talk less and work more, everybody would be better off.
10. No sane, normal, decent person could ever think of hurting a close friend or relative.

#### *Humanitarianism*

1. The death penalty is barbaric and should be abolished.
2. The dropping of the first atomic bomb on a Japanese city, killing thousands of innocent women and children, was morally wrong and incompatible with the standards of a civilized community.
3. Women have a moral right to complete equality with men in every sphere of life.
4. The Australian aborigines should immediately be granted full citizenship rights and complete social equality, including equal opportunities for education and employment, equal wages, etc.
5. Bodily punishment should not be applied to prison inmates, regardless of their behaviour.
6. Regardless of provocation, we should indulge in no brutality against enemy prisoners during time of war.
7. Under no circumstances whatever could a decision to make the first use of any nuclear or biological weapons be morally justified.
8. At the very least a quota of Asian migrants should be admitted to this country to demonstrate our goodwill towards our neighbours to the North.
9. Prostitutes should be regarded as victims of social circumstances who require assistance to become socially useful citizens.
10. Blood sports, such as fox hunting, greyhound racing with live hares, etc., are vicious and cruel and should be forbidden.

*Institutionalization*

1. The Church is necessary to establish and preserve concepts of right and wrong.
2. Every person needs to have the feeling of security given by a church.
3. For the vast majority of people, in order to live a truly religious life the Church or some such other organized religious body is an essential.
4. The aim of missionaries should be to establish church buildings where religious services and ceremonies can be conducted.

*Individualism*

1. A man ought to be guided by what his own experience tells him is right rather than by what any institution, such as the Church, tells him to do.
2. It is more important for an individual to understand the principles of his personal faith than to have a detailed knowledge of his own denomination.
3. Private devotions are more important in the religious life of a person than is attendance at public church services.
4. True Christianity is seen in the lives of individual men and women rather than in the activities of the Church.

*Beliefs about Christ*

1. (2) Jesus Christ was God the Son.
2. Jesus changed water into wine.
3. Jesus Christ was born of a Virgin.
4. Jesus walked upon the water while his disciples waited for him in their boat.

*Beliefs about God*

1. (1) There is a personal God.
2. (4) The world was created by God.
3. (7) There is a God who is all-powerful.
4. (8) There is a God who is altogether good.
5. God made man out of dust and breathed life into him.
6. (16) There is no God (personal or impersonal) (with scoring reversed).

*Other orthodox Christian beliefs*

1. (3) There are spiritual realities of some kind.
2. (5) There is a personal Devil.
3. (13) Evil is a reality.
4. (14) The spirits of human beings continue to exist after the death of their bodies.
5. (18) Attendance at church is a better way of spending Sunday than taking a walk in the country.

*General religious belief*

1. (6) Matter is the sole reality (with scoring reversed).
2. (9) There are such spiritual beings as angels.
3. (10) Jonah was swallowed by a great fish and afterwards emerged alive.
4. (20) Christianity is a better religion than Buddhism.
5. (21) The Bible is literally true in all its parts.
6. (23) There is a Hell in which the wicked will be everlastingly punished.
7. (25) Right will triumph.
8. There is no life after death (with scoring reversed).

*Opinion*

1. (15) Religion is the opium of the people.
2. (22) Man is, in some degree, responsible for his actions.
3. (27) Hardship strengthens character.
4. (32) Sex is evil.
5. (37) Members of the leisured class are supported by the 'surplus value' created by the workers.
6. (39) India has, on the whole, benefited from British rule.

*Fact*

1. (11) Man has been evolved from lower forms of life.
2. (17) The universe is expanding.
3. (19) Moses was the author of the first five books of the Bible.
4. (28) Mary, Queen of Scots, was beheaded between 1580 and 1590.
5. (30) Tigers are found in parts of China.
6. (31) Hornets live in nests under the ground.
7. (33) Light travels to us from the sun in less than one minute.
8. (34) Bacon was the author of the plays attributed to Shakespeare.
9. (35) Green is a primary colour.
10. (36) Sunlight is good for human health.
11. (38) Tariffs improve trade.
12. (40) The total national debt of Great Britain is more than a thousand million pounds.

*Miscellaneous*

1. (24) The spirits of persons who have died can sometimes communicate with the living.
2. (26) Belief in evolution is compatible with belief in a Creator.
3. (29) Everything is relative.
4. It makes no difference whether one is a Christian or not, so long as one has good will for others.
5. Salvation is only for Christian believers.
6. (12) There is an impersonal God.

*(Manuscript received 30 May 1961)*



# AN EXPERIMENTAL STUDY OF PIAGET'S THEORY OF THE DEVELOPMENT OF NUMBER IN CHILDREN\*

BY H. BLAIR HOOD

*Educational Psychologist, Cumberland*

In spite of the great weight of theory and speculation they contain, Piaget's books are all founded essentially on experimental work with children, and the purpose of this study was to repeat some of the experiments which Piaget has outlined in the book on number. Piaget has been criticized for not stating the number of subjects who took part in a given experiment from which he has drawn a specific conclusion: for not giving precisely the age range of the subjects, and for not relating their performance to mental age as well as to the chronological age. The aim of this study therefore was to determine whether (a) a group of normal English speaking children, and (b) a sample of mentally retarded children and adults, would show the same general trends in development of pre-number concepts as Piaget's subjects. The factors of mental and chronological age were also considered, as well as the relation of stage of development to ability in arithmetic.

## I. INTRODUCTION

In his foreword to *La Genèse de Numbre chez l'Enfant* Piaget (1941) indicates the book's position in the development of his thought. His early books were concerned with various verbal and conceptual sides of the child's thinking, covering such fields as language development, judgment and reasoning, ideas about the external world, physical causality and so on. Then come the studies of the origin of intelligence in children, from the moment of birth, until the end of the sensori-motor period of intelligence at about 2 years of age. This is followed by the book on number (Piaget, 1952) which is the first of a series of volumes devoted to discovering 'the mechanisms that determine thought'. All these aim at demonstrating how the various schemata of sensori-motor intelligence gradually cast off their sensori-motor dependence, and ascend to the level of relatively free conceptual thinking. Piaget does not consider sensori-motor intelligence to be thinking in the strict sense, but the schemata of this period of mental development do nevertheless become 'organized in operational systems on the plane of thought'. The purpose of these later volumes is therefore to trace the development of the operations which give rise to concepts of number, continuous quantities, space, time, etc., through the stage of intuitive pre-logic to that of concrete operations, and on to the final stage of formal operations.

When Piaget questions a child to determine the presence or absence of a given concept or logical operation, he usually employs two or more tests as a means of probing into the child's mode of thinking. We shall therefore use the word 'experiment' to mean the application of one or more tests to determine the stage of development of an operation, but the results tabulated will refer to the findings of the total experiment, not to individual tests. For instance, in chapter VII the problem of the

\* Based on a thesis submitted to Edinburgh University for the degree of Doctor of Philosophy.

additive composition of classes is studied by a number of tests in which the material used may be flowers, beads, or drawings of children; the important point is whether or not the various tests all indicate the same stage of development.

## II. SUBJECTS AND TESTING PROCEDURE

A sample of normal children was studied consisting of 61 boys and 65 girls, drawn from four schools in the north-west of England, and aged from 4.9 to 8.7. For the most part cultural backgrounds were satisfactory, with a country or small town bias. In all these schools good modern teaching methods were employed.

Also studied were 40 mentally subnormal subjects who fell into two classes as follows:

(a) *Educationally subnormal school children*. These are defined as 'pupils who, by reason of limited ability or other condition resulting in educational retardation, require some specialized form of education wholly or partly in substitution for the education normally given in ordinary schools'. The 23 subjects in the present study all had I.Q.'s below 75. All of them were boys attending a residential special school for such pupils. Their ages ranged from 10.3 to 15.7 years.

(b) *Mentally defective subjects*. These subjects were either (1) children of school age who had been classified as so defective in intelligence as to be unable to benefit from education in an ordinary school, or in a special school for educationally subnormal children: or (2) persons above school age who had been classified as mentally defective within the meaning of the Mental Deficiency Acts. All these subjects attended each day a training centre belonging to the local authority. There were 17 in all, their ages ranging from 9.8 to 41.0 years.

The Terman-Merrill scale 'L' (1937) was used to assess each child's mental age. Although much criticism has been levelled at the mental age concept and at the scale itself, it was felt that the scale could still differentiate quite usefully the social and intellectual levels of a group of children, giving a global view of their development; it was therefore used so that a comparison could be made between Piaget type tests and a widely accepted intelligence scale. Normally each child underwent three sessions of testing as follows: (1) The Terman-Merrill intelligence scale 'L' (1937); (2) the first four or five of the eight Piaget experiments selected for study; and (3) the remainder of the experiments. In a few cases a fourth session was required for checking previous results, or for assessing a child's arithmetical ability.

It should be stressed that the emphasis was on thoroughness of examination rather than on trying to test as many children as possible. Piaget's interrogatory method does not lend itself to high speed testing; it is by Piaget's choice a time-consuming clinical method in which every chance must be given to a child to express subtleties and nuances of thought, and at least so far as his present writings reveal, Piaget does not favour a standardized and uniform interrogation for all children. For this reason we did not force a standardized technique on to Piaget's work, although such a procedure may produce interesting and valuable results, as Lovell (1959) has reported. With very few exceptions, the children appeared to enjoy the session, and co-operation was excellent, even among the youngest.

When one comes to the crucial question in which the child is asked to make a judgment, e.g. 'Which has the most water?' 'Are there more flowers or more glasses?' etc., the children's responses may have to be accepted with great caution. Children seem to follow the rule 'When in doubt, repeat the last thing the man said'. It is therefore advisable to present each problem in several ways, if possible varying the form of the crucial question each time, and in this way one can be fairly sure in the end of the real character of the child's thinking.

## III. THE EXPERIMENTS

Eight of Piaget's experiments were selected for study. Each experiment aims at revealing the presence or absence of a pre-number concept, and at tracing the characteristic stages of its development. These will be described in bare detail only, without the nuances and theoretical ramifications which they contain in Piaget's book. The eight selected were:

(1) *The conservation of continuous quantities.* The child is shown, say, two large glass containers of similar dimensions and filled with an equal amount of liquid: he is allowed to satisfy himself that the amounts of liquid are the same. The liquid is poured from one container into four smaller ones, and the child is then questioned about the equality or otherwise of the two quantities as a result of this operation. Using a variety of containers the problem can be presented in several ways, and the stage of development clearly demonstrated.

(2) *The necessary equivalence of two sets of objects for which a one-to-one correspondence has been made.* The objects in the two sets are functionally related to one another, so that the correspondence will be to some extent provoked or suggested to the child, e.g. a set of glasses may be made to correspond to a set of bottles, or flowers to vases, etc. The equivalence of the two sets is recognized by the children while the one-to-one correspondence is present before them, but younger children lose the notion of equivalence when one set is closed up and the correspondence is no longer obvious.

(3) Given the following relation of equivalence  $x = y$ ,  $y = z$ , therefore,  $x = z$ , the question is whether the child can use this logical structure before he discovers *the mathematical notions of invariance of sets, seriations, etc.* Before he can arrive at the equation  $x = z$ , the child must presumably be at Stage 3 in the problem of the lasting equivalence of one-to-one correspondence of sets. The simplest technique is to have the child set up a one-to-one correspondence between jars and red flowers, put the red flowers to one side in a heap, make up a fresh correspondence between the same set of jars and blue flowers, and then question him as to the quantitative relation between the red and blue flowers.

(4) This experiment generalizes the function studied in the previous one to the point of *one-to-one correspondence between any number of sets.* For instance, since sets  $X$  and  $Y$  each correspond to  $Z$  in a one-to-one relationship,  $X + Y$  together will correspond to  $Z$  in a two-to-one relationship,  $W + X + Y$  in a three-to-one relationship, and so on. Piaget puts this experiment into a very complex setting, in which it appears as one of several experiments aimed at analysing the development of the multiplicative function from a number of different angles.

(5) *Establishing correspondence spontaneously* without having it suggested as in experiment 2, by constructing two sets of similar counters.

(6) *The construction of a series and the problem of ordinal correspondence.* The elements of a set differ in a manner which makes them capable of seriation, for instance a set of drawings of ten boys, each getting progressively smaller. The complementary set would be drawings of hoops corresponding in size to the boys, and the child would be required to put each set into serial order, and to make the two sets correspond.

(7) This experiment examines the child's ability to understand that *if class A*



consists of classes *B* and *C*, then  $B = A - C$ , etc. Flowers, beads, etc., may be used, or drawings of children. The child may for instance be shown a drawing consisting of a group of girls and a smaller group of boys, and the point is made that together they make up a group of children. The problem, which may then be presented in various ways, is to discover whether the child appreciates that although there are more girls than boys, there are more children than girls. Tests such as this may reveal the child's difficulty in thinking of the total class and at the same time one of its parts, largely because he is still thinking on the plane of perceptual intuition which is immediate and irreversible.

(8) The problem here is *the additive composition of numbers*, showing how the additive composition of parts into a whole gives rise to the same difficulty in numbers as we find in the inclusion of classes in a total class. A child may not understand that if he has four sweets in the morning and four in the afternoon, he is just as well off for sweets as his friend who has seven in the morning and one in the afternoon.

#### IV. RESULTS FOR NORMAL SUBJECTS

Table 1 shows the percentage of children for each of five age groups who are at one of three stages of development in each experiment. It also shows the average mental age for the group of children from whom each percentage was calculated. Although we cannot claim statistical precision for this table, which it might have had if a larger sample had been examined, the trends are fairly obvious. Pre-number concepts develop as the child gains in years, and most of those studied here are not fully developed until the child is between 6 and 7 years. It is not mere age, however, but mental stature which is primarily important. For instance, we may notice that in Experiment 2, the 64% in the under-five age group who are at Stage 2 have a mean mental age of 5.7, more than a year above the 28% of the same chronological age group who are at Stage 1, and whose mean mental age is only 4.3. This trend appears repeatedly throughout Table 1. If we regard the table as a guide to a practising teacher as to what to expect in the way of pre-number concepts from children in these age groups, it would surely suggest that chronological age grouping conceals the same diversity of pre-number development as of mental age development. For instance, seriation is largely not understood by around 40% of 5 to 6 year olds, but at the same time, almost as many have effective command of the concept, and between 20 and 30% in between these two stages have an unsteady and hesitant understanding of it. A teacher relying on chronological age grouping and making due allowance for exceptional cases, could feel reasonably sure that a large majority of her under-fives would have little appreciation of pre-number concepts, that those above seven could usually be counted on to approach full understanding, but that in the two year period from five to seven no assumption of any kind should be made.

Table 2 shows more positively the relation between stage of development and Terman-Merrill mental ages. Except for Experiments 2 and 8, 100% of the mentally under-fives are at Stage 1. Some concepts are beginning to emerge around mental ages 5-6 years but most children are still at Stage 1. Beyond the eight-year-old level, although 90-100% of responses are generally found to be at Stage 3, the evidence suggests that the concepts of seriation and class inclusion are still not fully under-

Table 1. *Stage of development and chronological age*

Experi- ment	Stage	Chronological age									
		4·1 to 5·0		5·1 to 6·0		6·1 to 7·0		7·1 to 8·0		8·1 to 9·0	
		%	Mean M.A.	%	Mean M.A.	%	Mean M.A.	%	Mean M.A.	%	Mean M.A.
1	1	86	5·2	94	5·11	47	6·2	—	—	—	—
	2	8	5·10	—	—	6	6·8	10	7·0	—	—
	3	4	5·10	6	7·8	47	7·3	90	8·1	100	8·10
2	1	28	4·3	6	4·6	—	—	—	—	—	—
	2	64	5·7	72	6·0	44	6·3	—	—	—	—
	3	8	6·0	22	6·7	56	7·1	100	7·11	100	8·10
3	1	84	5·4	69	5·7	33	6·0	—	—	—	—
	2	8	5·2	13	6·8	6	7·0	—	—	—	—
	3	8	6·0	18	6·7	61	7·4	100	7·11	100	8·10
4	1	100	5·3	78	5·9	47	6·3	—	—	—	—
	2	—	—	14	6·4	24	6·10	—	—	—	—
	3	—	—	8	7·8	29	7·8	100	7·11	100	8·10
5	1	72	5·1	40	5·4	21	6·0	7	7·0	—	—
	2	24	5·8	26	6·2	23	6·7	—	—	—	—
	3	4	5·10	34	6·8	56	7·4	93	8·1	100	8·10
6	1	96	5·3	85	5·10	52	6·4	36	7·4	25	9·2
	2	4	5·10	15	6·6	42	7·4	28	7·9	—	—
	3	—	—	—	—	6	8·6	36	9·0	75	8·9
7	1	96	5·2	92	5·11	85	6·11	45	7·7	25	8·0
	2	—	—	—	—	15	7·2	—	—	—	—
	3	4	5·10	8	6·8	—	—	55	8·5	75	9·2
8	1	72	5·1	78	5·10	44	6·5	—	—	—	—
	2	4	5·10	—	—	6	7·2	10	7·2	—	—
	3	24	5·7	22	6·9	50	7·3	90	8·1	100	8·10
<i>n</i> for each age group		23		32		34		28		9	

Table 2. *Stage of development and mental age*

Experi- ment	Mental age														
	Below 5·0			5·1-6·0			6·1-7·0			7·1-8·0			8·1+		
	Stage			Stage			Stage			Stage			Stage		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
	( <i>%</i> )			( <i>%</i> )			( <i>%</i> )			( <i>%</i> )			( <i>%</i> )		
1	100	—	—	71	12	17	64	6	30	9	9	82	—	—	100
2	48	52	—	—	85	15	—	44	56	—	—	100	—	—	100
3	100	—	—	60	11	29	44	11	45	—	—	100	—	—	100
4	100	—	—	100	—	—	37	18	45	—	24	76	—	8	92
5	100	—	—	50	29	21	10	20	70	—	6	94	—	—	100
6	100	—	—	100	—	—	45	35	20	18	49	33	23	25	52
7	100	—	—	87	—	13	84	5	11	54	20	26	42	—	58
8	70	6	24	83	—	17	62	8	30	9	9	82	—	—	100
<i>n</i> for M.A. groups	17			38			31			26			14		

stood by 50 %. In contrast to the chronological age grouping, we find that it is in the mental age band 6-8 years inclusive that the change from Stage 1 to Stage 3 responses is most actively taking place, but this trend, of course, may apply only to evidence based on Terman-Merrill age levels.

The relation between mental age and stage of development is shown collectively for all the experiments in Table 3. Here we have half-year mental age groups, and the table records the number of responses at a given stage of development which were derived from all the experiments. Thus at age 4.0-4.5, the five children in this age group gave Stage 1 responses in all the experiments, making forty such responses in all. It will be seen that the results of this table were obtained from 77 children only, and unfortunately, owing to practical considerations, it was not found feasible to extend the table to include the results of the whole sample. By the age of about 7 years, 50 % of the responses are at Stage 3. It is also again clear from this table that, for Terman-Merrill mental ages, the great period for the development of pre-number concepts is between 6.0 and 8.0, when the proportion of Stage 3 responses increased by nearly 70 %.

The percentage is 50 at about age 7.1. The 'expected' Stage 3 responses were fitted by an approximate probit analysis.\*

Table 3. *Relationship between M.A. and frequency of responses at each stage of development*

M.A. group	N	No. of responses at Stages			Expected %	
		1	2	3	at Stage 3	for following ages
4.0-4.5	5	40	0	0	1	4.6
4.6-4.11	6	46	2	0	2	5.0
5.0-5.5	7	47	9	0	6	5.6
5.6-5.11	11	65	15	8	14	6.0
6.0-6.5	14	74	18	20	28	6.6
6.6-6.11	7	21	15	20	46	7.0
7.0-7.5	10	18	19	43	65	7.6
7.6-7.11	3	0	2	22	81	8.0
8.0-8.5	6	5	1	42	91	8.6
8.6-8.11	5	2	5	33	97	9.0
9.0-9.5	3	0	0	24	99	9.6

#### V. PIAGET'S PRE-NUMBER CONCEPTS AND ABILITY IN ARITHMETIC

The question now is whether there is any relationship between the presence or absence of pre-number concepts and ability in arithmetic. There is, of course, no scale of arithmetical ability which will correlate with a given number of passes in the Piaget tests so that it would be possible to predict the one from the other. In the first place, as we have seen, in the present state of our knowledge the concepts assessed by the experiments make their appearance over a broad time belt from at least 6.0 to 8.0 years; nor has it been possible yet to indicate a normal and expected pattern of passes and failures in the Piaget tests for any specific mental age, although preliminary work on such a scale has been reported by Vinh Bang (1957). Secondly, a valid scale of arithmetical ability is very difficult to obtain with young children.

\* Dr D. N. Lawley kindly gave advice on the use of this statistic.



Even among the four schools of this investigation, there was a good deal of variation in teaching methods, time devoted to arithmetic, age of entry to school, and of course such factors as absence from school, change of teacher, etc. Piaget would never intend his experiments to be measures of how much arithmetic a child should be able to do. What they try to do is to indicate aspects of mental structure, so that we may know if a child has developed mentally to the point of having those concepts, which are the intellectual foundations of arithmetic, and without which a complete understanding of arithmetic is impossible. We may, however, anticipate that some children who fail in all the Piaget tests may nevertheless be drilled so expertly in mechanical arithmetic that they appear to have real skill in it, whereas their lack of ability becomes evident when they are faced with simple problems, or more advanced mechanical work.

Class teachers were asked to review each child's work in arithmetic, and to place the child in one of five arithmetic grades.

*Grade 1.* No number ability of any kind: child may be able to count, but he cannot pick out five or more objects from a group, or do a simple sum even with counters.

*Grade 2.* He can pick out five or more objects from a group.

*Grade 3.* Child can do simple addition and subtraction with or without counters, orally or in writing.

*Grade 4.* He can do simple problems presented verbally or in writing, and also will have the skills of Grade 2; he can in short use number with apparently some understanding.

*Grade 5.* Any combination of number skills beyond this point.

Teachers did not find it easy to fit every child clearly into one or other of these grades, and many of them asked for a finer discriminating scale. The above scale, however, was thought to be adequate, since its purpose was merely to pinpoint any child who might be able to use number with understanding, and yet be predominantly lacking in the pre-number concepts. By 'using number with understanding', perhaps better expressed as 'using numbers successfully', we mean the ability to attend to the numerical elements in a number problem, to relate these elements correctly to one another, and to resolve the correct numerical solution to the problem from these relationships. The process must be distinguished from that of 'understanding number'. For Piaget 'number' as such is understood only when the various concepts of conservation, equivalence, equalization of differences, and so on, can be attached to a group of interchangeable units, and this is assessed for a given child not by arithmetical problems, but by the techniques outlined in his book. Usually it would be denied that a child can use number successfully before he can understand numbers, but our evidence does not support this view. Arithmetic logically entails the concepts of conservation, equivalence, and the rest, and would not make sense without them, but psychologically a degree of arithmetic is possible for children predominantly or entirely at Stage 1. A child may be trained, not only in mechanical processes, but in problem work, to act as if he understood number. Methods of solving a problem may be skilfully taught, and are, moreover, facilitated when the child solves the problem by acting on material before him. The material itself helps to keep his thinking in the channels of conservation, equivalence, etc., and the presence or absence of the concepts themselves does not constitute for him an element in the problem.

Table 4, based on teachers' estimates of arithmetic grade, shows that at least an elementary degree of arithmetic may be expected from many children who have still not acquired Piaget's pre-number concepts, but at the same time, we find no case of a predominantly Stage 1 child being in Grade 5; although under very favourable circumstances such a case may be found, it seems probable that as a general rule more advanced work than the level of Grade 4 will increasingly demand these concepts. One should not expect a child of predominantly Stage 1 responses to go far in arithmetic, and there is a strong case for regarding presence or absence of the concepts as the main factor in weighing up a child's arithmetic readiness on the intellectual plane. On quite other grounds, however, a proportion of Stage 3 children will be found to be at too low a standard of arithmetic, their backwardness arising from the numerous vicissitudes of school learning and not from mental incapacity.

Table 4. *Arithmetic grade and stage of development*

Arithmetic grade	Percent totally or predominantly at Stage	
	1 and 2	3
1	37	0
2	23	12
3	25	13
4	15	13
5	0	62

Much of the validity of Piaget's pre-number theory hangs on the soundness or otherwise of his experimental technique, and we therefore find that he anticipates possible objections to his method of study. 'It might be argued', he says, in considering the responses of children at Stage 1, 'that the mistakes are due to lack of understanding of the words used. May it not be that the child does recognize that the number of bottles and glasses remains the same when one set is grouped together, and that when he says "there are more" he is merely expressing the idea that the shape of the set has changed and the space it occupies is greater.' This point presented itself to the writer frequently with force. For instance, many children who gave Stage 1 type answers ('there are more here', 'this is bigger', etc.) were asked a further question, 'Do you mean that you have now a bigger *number* of things here than here?' or 'If you counted the line of flowers would you now count more than if you counted the heap?'. A few children gave a clear answer to this, showing that they knew the number of items in each set to be the same; for instance, 'It's only the way they look that's different', and they were then passed as Stage 3. Many of them, however, agreed that the number was unchanged, yet when the tests were repeated they persisted in saying that one set was more or bigger than the other. It became a question then of to what extent one should insist on them thinking only of the number in each set by, as it were, helping them by verbal suggestion to set aside in their minds irrelevant perceptual factors. This procedure would have given a somewhat earlier Stage 3 development, but it would also have meant that the examiner was doing a good deal of the child's thinking for him. The mere fact that a child tends to look at the quantitative problem set to him, not solely in analytic terms of 'how many times', but in global, perceptual terms, is evidence surely of immature numerical

thinking, however much we may feel that he 'really knows all the time that there are still six in each set'. The fact is that the child, although he may or may not know that the sets are still six and six, cannot prevent himself lumping in with the six items the space that lies in between, or the size of the individual items in one set as compared with those in the other. Partly to throw some light on the point that the child's difficulty may be largely a linguistic one, implying that he merely does not understand what the examiner is asking for, the writer in the course of Experiment 7 from time to time took a clear Stage 1 child, and carefully explained to him just why his answer was wrong, and what the correct answer should have been, using the three types of apparatus. Nineteen children were 'coached' in this way and were retested after an interval of 3-4 days. In every case the child on being retested was still at Stage 1. It seemed evident too that when a child was questioned immediately after the working session, he could usually give the correct response, but mainly because he was repeating to the examiner what he had in effect been told to say, and there was little or no evidence of temporary insight into the problem.

Piaget answers his critics by stressing the confidence and clarity of mind which the child displays when he does reach Stage 3. He also makes the important point that even if the child can count the items of the sets correctly, there is no proof that this 'verbal enumeration expresses a better quantification, from the child's point of view, than the space occupied'. Correspondence between numerals and objects at this level may be purely verbal—just as the child can make the glasses correspond to the bottles, so he can make the names of the first six numerals correspond to the six glasses, etc.

While it is certainly true that there is an element of linguistic confusion among Stage 1 children because their understanding of the words in question is incomplete and limited, it is the eventual appearance of the concept and operation, the internalized action becoming reversible, that completes the verbal understanding and removes the limitations. With notions such as permanence and equivalence now a part of his mental equipment, the child modifies and re-directs his use of expressions such as 'which is bigger, where are there more . . . , etc.'. Before this point is reached, one may try to teach children the correct quantitative use of such expressions, so that, as some would have it, 'we implant a sense of number in their minds'. With young children such teaching would bear no fruit, but from approximately 5 or 6 years onwards, it should be possible to help a child on more quickly to Stage 3, using such didactic material as, say, the Cuisenaire rods, and certainly the number teaching of young children should be directed to this end. Basically, however, Stage 3 is reached by the normal process of intellectual maturation.

## VI. RESPONSES OF MENTALLY SUBNORMAL SUBJECTS

In his theory of the growing mind of the child, Piaget gives us a picture of the evolution of a mental structure which, while retaining always the unity of its inter-related parts, increases steadily in size and complexity until the period of formal operations is completed, system following on system and each integrated with the one which preceded it. The child's pre-number concepts are a fragment of this growing system of operations. They emerge over the age range of from 6 to 10 years



approximately as concrete operations, and along with other concrete operations, they are the base on which the final structure of formal operations is built. Knowing therefore their time and position in this procession of mental events among normally developing children, we may now ask whether their development is affected by the handicap of general mental subnormality. For instance, a mentally subnormal boy of 19 years may have an intelligence test mental age of 6 years. Will his thought processes still be stumbling at a pre-concrete operational level, or shall we find that his thinking, as expressed in operations, has matured along more normal lines because of his contacts and experiences in the society of his home, school or training centre? If we find that he has in effect 'learned' the answers to the pre-number problems, our conclusion must be that the successful solution of such problems is not so firmly, if at all, pinned down in a system of operations. Persuasive as Piaget's arguments are, one wonders from time to time—particularly perhaps while in the process of testing children—whether, say, the child who solves the glasses of water problem does so primarily because of his home and school experience, and only in a secondary sense because of a scheme of mental operations. The reaction of chronologically older subjects, however, who have mental ages numerically comparable with the young child, may help to reinforce one or other point of view.

At least two studies on these lines have already been undertaken. Inhelder (1943), sets out with the full intention of demonstrating how Piaget's theory of the system of reversible operations could help materially in the diagnosis of various grades of mental deficiency. More recently Woodward (1959) has applied Piaget's observations on sensori-motor intelligence and the concept of permanent objects in the normal child's development, to a sample of mentally defective children. She concludes that 'sensori-motor activities of severe mental defectives could be classified into the sensori-motor types distinguished by Piaget'.

In their manner of responding to the examination, the mentally retarded group tended to differ from the normal children in the following ways: (1) The retarded subjects took on the average approximately 10 min. longer to complete each session of the examination. They moved more slowly, both mentally and physically, and one very seldom found the speedy, alert response which even some normal 5 year olds could display. (2) This slowness was due in part to their poor receptiveness to language. One had to speak to them more slowly and deliberately, with repetition being more frequently necessary. Interpretation of their comments and observations, although they were seldom incoherent, was no easy task even when a response was essentially correct. Frequently a long verbal reply would be given, with the correct response buried somewhere among the irrelevancies. (3) Even among those with higher mental ages, distractibility was common, taking the form mainly of wanting to play with the test material, or breaking in with quite irrelevant questions or observations. These points were of course more in evidence among the mentally defective than the educationally subnormal subjects. Neither group showed any evidence of ill-will or resistance to the examinations, and they gave the impression that as long as one did not hurry them, they would work to the limit of their capacity.

The responses obtained from these subjects showed that as with young children of normal intelligence, Piaget's pre-number concepts tend to mature as the subject's

all-round mental capacity increases. It is suggested that general socialization, the effects of schooling, and the training of an occupation centre might have provoked the maturation of these concepts among the chronologically older subjects whose mental development on a standard intelligence scale was still theoretically that of a young child. For instance, to take an extreme case, one subject was a man of 40 who was attending the occupation centre as an interim measure until more suitable arrangements could be made for him (he was well beyond the usual age limit for the centre). Because of his tendency to down tools when he felt like it, and his limited sense of responsibility in most social matters, he was never a satisfactory worker even in an unskilled job. Nevertheless he had 'knocked about a bit', could handle small quantities of money needed for cigarettes, buses, etc., and in casual conversation did not give the impression of being so low on the mental age scale. But for all his worldly experiences, he remained solidly at Stage 1 in all the experiments, with no hint at all of even approaching Stage 2 at any point. He also had the normal 5 year old's frequent difficulty in coping with such a phrase as 'Make it the same as . . . , which one is bigger . . . , have I got more . . . , etc.'. Several other subjects showed a similar pattern of responses.

Table 5. Showing 'expected' percentage of Stage 3 responses for each age group of mentally retarded subjects

Age	5.0	5.6	6.0	6.6	7.0	7.6	8.0	8.6	9.0	9.6	10.0	10.6	11.0	11.6
Expected % Stage 3	1	3	5	9	15	24	34	46	59	70	80	87	93	97

*Italicized figures extrapolated.*

Turning to a higher level on the mental age scale, we find that from about  $8\frac{1}{2}$  years upwards there is a preponderance of Stage 3 responses. In nearly every test these subjects gave the quick and confident responses of normal children, including often the same element of surprise and amusement at the obvious simplicity of the problems. The association between these subjects' mental retardation and their immaturity in pre-number concepts can be studied by obtaining for each subject the number of responses he gave at each stage of development from all the experiments taken together. From this the 'expected' percentage of Stage 3 responses at each mental age level for the mentally retarded subjects can be calculated on the same lines as the last two columns of Table 3. The complete table of individual results from which Table 5 was constructed is not reproduced, but some of the results stand out as anomalous cases when compared with the general pattern of responses shown in Table 3. The following three subjects, for instance, show an overloading of Stage 1 responses which one would not expect from their Terman-Merrill mental age:

Subject	M.A.	Total of responses at Stage		
		1	2	3
1	7.8	8	0	0
2	8.0	4	1	3
3	9.6	5	1	2

approximately as concrete operations, and along with other concrete operations, they are the base on which the final structure of formal operations is built. Knowing therefore their time and position in this procession of mental events among normally developing children, we may now ask whether their development is affected by the handicap of general mental subnormality. For instance, a mentally subnormal boy of 19 years may have an intelligence test mental age of 6 years. Will his thought processes still be stumbling at a pre-concrete operational level, or shall we find that his thinking, as expressed in operations, has matured along more normal lines because of his contacts and experiences in the society of his home, school or training centre? If we find that he has in effect 'learned' the answers to the pre-number problems, our conclusion must be that the successful solution of such problems is not so firmly, if at all, pinned down in a system of operations. Persuasive as Piaget's arguments are, one wonders from time to time—particularly perhaps while in the process of testing children—whether, say, the child who solves the glasses of water problem does so primarily because of his home and school experience, and only in a secondary sense because of a scheme of mental operations. The reaction of chronologically older subjects, however, who have mental ages numerically comparable with the young child, may help to reinforce one or other point of view.

At least two studies on these lines have already been undertaken. Inhelder (1943), sets out with the full intention of demonstrating how Piaget's theory of the system of reversible operations could help materially in the diagnosis of various grades of mental deficiency. More recently Woodward (1959) has applied Piaget's observations on sensori-motor intelligence and the concept of permanent objects in the normal child's development, to a sample of mentally defective children. She concludes that 'sensori-motor activities of severe mental defectives could be classified into the sensori-motor types distinguished by Piaget'.

In their manner of responding to the examination, the mentally retarded group tended to differ from the normal children in the following ways: (1) The retarded subjects took on the average approximately 10 min. longer to complete each session of the examination. They moved more slowly, both mentally and physically, and one very seldom found the speedy, alert response which even some normal 5 year olds could display. (2) This slowness was due in part to their poor receptiveness to language. One had to speak to them more slowly and deliberately, with repetition being more frequently necessary. Interpretation of their comments and observations, although they were seldom incoherent, was no easy task even when a response was essentially correct. Frequently a long verbal reply would be given, with the correct response buried somewhere among the irrelevancies. (3) Even among those with higher mental ages, distractibility was common, taking the form mainly of wanting to play with the test material, or breaking in with quite irrelevant questions or observations. These points were of course more in evidence among the mentally defective than the educationally subnormal subjects. Neither group showed any evidence of ill-will or resistance to the examinations, and they gave the impression that as long as one did not hurry them, they would work to the limit of their capacity.

The responses obtained from these subjects showed that as with young children of normal intelligence, Piaget's pre-number concepts tend to mature as the subject's



all-round mental capacity increases. It is suggested that general socialization, the effects of schooling, and the training of an occupation centre might have provoked the maturation of these concepts among the chronologically older subjects whose mental development on a standard intelligence scale was still theoretically that of a young child. For instance, to take an extreme case, one subject was a man of 40 who was attending the occupation centre as an interim measure until more suitable arrangements could be made for him (he was well beyond the usual age limit for the centre). Because of his tendency to down tools when he felt like it, and his limited sense of responsibility in most social matters, he was never a satisfactory worker even in an unskilled job. Nevertheless he had 'knocked about a bit', could handle small quantities of money needed for cigarettes, buses, etc., and in casual conversation did not give the impression of being so low on the mental age scale. But for all his worldly experiences, he remained solidly at Stage 1 in all the experiments, with no hint at all of even approaching Stage 2 at any point. He also had the normal 5 year old's frequent difficulty in coping with such a phrase as 'Make it the same as . . . , which one is bigger . . . , have I got more . . . , etc.'. Several other subjects showed a similar pattern of responses.

Table 5. Showing 'expected' percentage of Stage 3 responses for each age group of mentally retarded subjects

Age	5.0	5.6	6.0	6.6	7.0	7.6	8.0	8.6	9.0	9.6	10.0	10.6	11.0	11.6
Expected % Stage 3	1	3	5	9	15	24	34	46	59	70	80	87	93	97

Italicized figures extrapolated.

Turning to a higher level on the mental age scale, we find that from about 8½ years upwards there is a preponderance of Stage 3 responses. In nearly every test these subjects gave the quick and confident responses of normal children, including often the same element of surprise and amusement at the obvious simplicity of the problems. The association between these subjects' mental retardation and their immaturity in pre-number concepts can be studied by obtaining for each subject the number of responses he gave at each stage of development from all the experiments taken together. From this the 'expected' percentage of Stage 3 responses at each mental age level for the mentally retarded subjects can be calculated on the same lines as the last two columns of Table 3. The complete table of individual results from which Table 5 was constructed is not reproduced, but some of the results stand out as anomalous cases when compared with the general pattern of responses shown in Table 3. The following three subjects, for instance, show an overloading of Stage 1 responses which one would not expect from their Terman-Merrill mental age:

Subject	M.A.	Total of responses at Stage		
		1	2	3
1	7.8	8	0	0
2	8.0	4	1	3
3	9.6	5	1	2

Other cases show the same trend, if not so strongly. Thus although there is clear relation between mental age and stages of development, the former is not by any means an invariable guide to the presence or absence of pre-number concepts. It is not surprising to find, however, that when the responses of these three subjects in the Terman-Merrill scale were examined it was found that they tended to score well on items which were relatively free of logical content and which relied on such factors as memory and social training and experience. On the other hand, these subjects were obviously weak in items calling for the identification of similarities and differences, and in other conceptual problems. The supervisor of the training centre commented that while superficially these three children might behave like normal and reasonably sensible 8-9 year olds, she recognized that underneath their general retardation there was a grosser and more intellectually crippling handicap which showed itself only in special circumstances. Much the same opinion was expressed of other subjects. From Table 5 we find that mentally retarded subjects reach 50 % Stage 3 responses at about mental age 8.8, compared with 7.1 among normal children. Their rate of development, as one would expect, is also much slower; between 6 and 8 years their proportion of Stage 3 responses increases by only 25 % compared with nearly 70 % among normal children.

The expected percentages were again determined by an approximate probit analysis. The percentage is 50 at about age 8.8. The results have only limited reliability since they are based on the results of only forty subjects.

An interesting point came to light in the relation of stages of development in pre-number concepts and mental ages among the educationally subnormal pupils, which had a bearing on their ability in arithmetic. Individual reports on each pupil's ability in arithmetic had been obtained from the class teacher. These reports were easily given a threefold classification as follows:

- (a) Pupils with virtually no arithmetical ability: some of them could not count, others might be able to do a simple adding sum using counters, but their capacity to do this was shaky and unreliable.
- (b) Pupils who could do simple addition and subtraction sums, but whose progress came to a stop with the introduction of carrying and borrowing.
- (c) Pupils who, although generally about 18 months to 2 years behind their mental ages in arithmetic, showed definite understanding of and confidence in what they were doing.

There were 8, 7 and 8 pupils respectively in these categories.

One noticed that the pupils who failed most heavily on the pre-number concept tests were all very weak in arithmetic. But most of those who passed the Piaget tests successfully were making reasonable progress in number work. There were, however, a few exceptions to this where there was poor arithmetic in spite of good responses in the Piaget tests. There was no instance of success in arithmetic and failure in the Piaget tests being found in the same child. Although the teachers of these boys expected them to be backward in arithmetic compared with normal children of the same chronological age, they did nevertheless expect a boy whose mental age was 7 or more to show some degree of elementary understanding of arithmetic. They had access to test results and knew their pupils' mental ages, and expected to get the type of responses characteristic of the normal 7 year old. A number of boys with mental ages from 7 to 8 years were therefore presented to the

writer as special problems, particularly since they were found to be considerably more successful in reading than in arithmetic, and therefore 'were not by any means fools'. The results of the Piaget tests strongly suggested that their failure in arithmetic was due to the absence of pre-number concepts, and that in fact, in spite of their Terman-Merrill mental ages, which were checked by the writer at the time, they were mentally not ready for arithmetic. They were still at Piaget's intuitive stage of thinking, and had not yet reached the stage of concrete operations.

A study of a few individual case notes, which space will not allow to be quoted here, would illustrate the profound difficulty defectives may have in handling such words and expressions as *more, most, bigger than, the same as, equal to, like each other*, etc. These expressions imply a quantitative comparison, and it is in making the comparison that the subjects' thinking breaks down. They must be able to attend to two objects, say jar *A* and jar *B*, or a set of red flowers and a set of blue flowers. The two objects must be linked by a concept of quantity, e.g. size, amount, number. The concept becomes active or operational when the mind can pass in attention from one subject to the other, and then back to the first. Unless it has this reversible mobility along with a distinct orientation towards quantity, it cannot be truly operational, and no comparison of quantity can be made. Subjects at Stage 1 can make only a uni-dimensional comparison, i.e. jars *A* and *B* must be identical, with the only difference being level of water: the set of flowers must be spaced similarly to the blue, the only difference being total length. When, however, two-dimensional differences are introduced, e.g. jar *A* different in shape from *B*, or the blue flowers different in density as well as length from the red, a second element of the concept of quantity must be introduced to the first schema, namely the conservation of quantity—as Piaget has elaborated. The responses of the normal 5 year olds in this investigation suggest that although they may not have reached the level of a schema having conservation of quantity, they have reached the simple form of quantitative comparison referred to above. On the other hand many defective subjects whose test mental ages are 5–6 years have not yet reached even this simple schema, and quantitative comparisons even at a uni-dimensional level, are therefore beyond them. Faced with two objects of differing size, the percept of each remains a relatively passive unit in his thinking, each percept standing alone and unrelated to the other. Only mobility can bring about the unifying of the two percepts into one larger mental unit, or elementary structure, thus introducing into thought the possibility of making quantitative comparisons.

The writer's thanks are due to Dr Mary Collins who advised him on many points during the various stages of this study. Thanks are also due to the educational and medical authorities of Cumberland and the City of Carlisle in whose schools and training centres these subjects were examined.



## REFERENCES

- INHELDER, B. (1943). *Le Diagnostic de Raisonnement chez les Debiles Mentaux*. Neuchatel: Delachaux et Niestle.
- LOVELL, K. (1959). A follow-up study of some aspects of the work of Piaget and Inhelder on the child's conception of space. *Brit. J. Educ. Psychol.* **29**, 104-17.
- PIAGET, J. (1941). *La Genèse du Nombre chez l'Enfant*. Neuchatel: Delachaux et Niestle.
- PIAGET, J. (1952). *The Child's Conception of Number*. London: Routledge and Kegan Paul.
- VINH BANG (1957). Elaboration d'une Échelle de Developpement du Raisonnement. *Proceedings of the Fifteenth International Congress of Psychology*, Brussels, 1957.
- WOODWARD, M. (1959). The behaviour of idiots interpreted by Piaget's theory of sensori-motor development. *Brit. J. Educ. Psychol.* **29**, 60-71.

(*Manuscript received 6 June 1961*)

## GROUP INFLUENCE ON THE PERCEPTION OF AMBIGUOUS STIMULI

By A. E. M. SEABORNE

*Social and Psychological Research Centre, Brunel College*

An attempt was made to relate the judgements made by subjects in a social pressure situation to the ambiguity of the stimulus being judged. Ambiguity values were assigned to the various stimuli on the basis of responses made by subjects to the stimuli when not exposed to social influence.

It was found that conformity was not related to the ambiguity of the stimulus although individual certainty of correctness varied inversely and response time varied directly with ambiguity. It seems possible that this can be accounted for by assuming that the subject's assessment of the difficulty of the judging task was used to rate the value of the group's opinion so that where it seemed that the possibility of arriving at a correct answer was less, the group's influence was less effective.

The results indicate that conformity can only be expected to vary directly with ambiguity where the social influence is seen to derive from a source external to the immediate stimulus situation.

### INTRODUCTION

It seems to be accepted that ambiguity of stimulus is a factor accentuating the accepting of social influence in suggestion situations, e.g. Coffin (1941).

This experiment is concerned with the question whether ambiguity of a display, as defined in a particular way, increases conformity to group pressures.

The basic situation used to test this idea was one in which the subject had a choice of correct answers. Where either of two answers can be correct, the probability of either being seen as correct is less than where only one is correct. If two alternative stimulus situations can be set up, one where there is only one correct answer and one where there are two correct answers, and if the situations are such that the combined chances of getting a correct answer in the double-answer case, so that the difficulties in chance of getting a correct answer in the single-answer case, then the subject responding to the two-situation are equal, then we would predict that the subject responding to the two-answer stimulus in the presence of group pressures to give an incorrect answer would be more likely to conform than in the single-answer case.

Thus, although the double-answer case would be no more difficult, in that the probability of a veridical perception would be as high in the individual situation, in the group situation the subject would be more likely to give a response which is incorrect.

The experimental situation used was a modification of the arrangement used by Asch (1956). The subjects were presented with two cards, on one of which was the standard line and on the other several comparison lines, and the explicit task was to judge which of the comparison lines was equal to the standard. All except one of the subjects in the group were told which response to make: the remaining subject was the critical subject whose responses were being studied.

It was decided to use four levels of ambiguity as described, so comparison cards on which there were one, two, three, and four correct answers were made. Difficulty was also varied in that the difference in length between the comparison lines equal to

the standard and the other comparison lines was varied. First, it was necessary to design a set of cards in which the difficulty of the two, three and four correct answer cards was equal to the difficulty of the one answer card. It was decided to use a range of difficulty extending from about the same as Asch's situation to a stage where only chance correct results would be achieved.

Thus for two purposes the cards had to be tested out on a number of people individually: first, to ensure that the objectively correct answers were equal in the two, three and four correct answer cases to the single correct answer case, and, second, to arrive at an ambiguity value based empirically on the actual distribution of responses. The proposed definition of ambiguity is not in terms of number of possible correct answers but in terms of the actual number of lines seen as correct, by a sample of individuals.

#### *Display design experiment*

Forty comparison cards were therefore designed, and tested in an individual experimental situation.

The properties of the forty comparison cards can be seen in Table 1. The standard line was 4 in. long in all cases. The lengths of the lines on card no. 1 were  $3\frac{1}{16}$ , 4 and  $4\frac{1}{8}$  in. Card no. 2 had six lines, two each of these three lengths. Cards 3 and 4 had nine and twelve lines which were three and four sets of the three lengths. The lines on card 5 were  $3\frac{1}{8}$ , 4 and  $4\frac{1}{4}$  in. and on card 6 there were two sets of these.

Table 1. *Cards used in display design experiment*

Size difference (in.)	Standard centre. Number of lines				Standard shortest. Number of lines			
	3	6	9	12	3	6	9	12
	Card no.							
$\frac{1}{16}$	1	2	3	4	21	22	23	24
$\frac{1}{8}$	5	6	7	8	25	26	27	28
$\frac{3}{8}$	9	10	11	12	29	30	31	32
$\frac{1}{2}$	13	14	15	16	33	34	35	36
$\frac{3}{4}$	17	18	19	20	37	38	39	40

It can be seen on Table 1 that there were five levels of difficulty in that there were five sets of cards on which the lengths of the lines differed among themselves in steps of  $\frac{1}{16}$ ,  $\frac{1}{8}$ ,  $\frac{3}{8}$ ,  $\frac{1}{2}$  and  $\frac{3}{4}$  in., respectively.

There were two sets of twenty cards each. The first set from cards 1 to 20 had the comparison line equal to the standard (the 4 in. line) at the centre of the distribution of three lengths and the other set from cards 21 to 40 had the standard length as the shortest line. So the lines on card 21 were 4,  $4\frac{1}{8}$  and  $4\frac{1}{4}$  in. long. (This was done to avoid a situation in which the task was to choose the central length rather than the length equal to the standard.) The cards, therefore, varied in two major dimensions, level of difficulty represented by size difference and ambiguity represented by number of lines, and also according to whether the correct line was at the centre or the bottom of the distribution.

The relative orders of lines on the cards from left to right were varied randomly. The cards were  $12 \times 8$  in. and the lines were drawn parallel to the shorter side. They were  $\frac{3}{16}$  in. thick and were spaced at  $\frac{3}{4}$  in. intervals symmetrically about the mid-line of the longer side of the card. Their relative base lines were varied randomly except that no line started or ended less than 1 in. from the edge of the card. A letter was printed under each line for identification.

An order of presentation of the forty cards was then designed so that all size differences were represented within each set of five and all numbers of lines within each set of four. The experimental procedure was as follows.

The cards were presented once through to twenty individual subjects. The cards were ordered



in two subsets of twenty each, subsets A and B. In order to minimize sequence and position effects the subjects were divided into four groups. With the subjects of the first group the standard card was on the left of the stimulus card in the frame and subset A was presented first, with the second group the standard was on the left, but subset B was presented first: and so on. The subject sat 12 ft. from the stimulus frame and was told that he would be shown a succession of pairs of cards one of which would have one line on it and the other several lines. His task was to decide which of the several lines was equal to the single line and to give as his response the letter under the line he thought equal.

The stimulus frame was simply a device whereby two cards could be held side by side at about the eye-level of subjects seated 12 ft. away. The frame hinged backwards and when the subject had responded the experimenter dropped the frame backwards, changed the cards and replaced it.

The results in terms of number of correct responses are shown in Table 2. As can be seen the proportion of correct responses increased as the size difference between the lines increased and the number of correct responses did not differ significantly between the different number of lines.

Table 2. *Number of correct answers in display design experiment*

Size difference (in.)	Standard centre. Number of lines				Standard shortest. Number of lines			
	3	6	9	12	3	6	9	12
$\frac{1}{16}$	3	6	5	3	5	4	4	5
$\frac{1}{8}$	10	7	9	9	5	8	14	11
$\frac{1}{4}$	13	7	14	13	10	11	14	12
$\frac{3}{8}$	14	14	17	16	15	16	16	17
$\frac{1}{2}$	17	19	20	19	19	19	18	18

*Total correct answers by size difference*

$\frac{1}{16}$ in.	35	} out of a possible	160 = 21.8 %
$\frac{1}{8}$ in.	73		160 = 45.6 %
$\frac{1}{4}$ in.	94		160 = 58.7 %
$\frac{3}{8}$ in.	125		160 = 78.1 %
$\frac{1}{2}$ in.	149		160 = 93.1 %

*Total correct answers by number of lines*

3	} lines	111	} out of a possible	200 = 55.5 %
6		111		200 = 55.5 %
9		131		200 = 65.5 %
12		123		200 = 61.5 %

Total percentage correct = 59.5 %.

It was decided that forty cards were too many to use in a group experiment, as this made the task long and very boring for the subject, so half of the cards were selected. The principles of selection used were that approximately half of the chosen twenty should be cards having the line equal to the standard at the centre of the distribution of lengths, and half at the bottom of the distribution, and that the cards should be 'typical' of their sets, i.e. that objective difficulty should increase as the size differences decreased and that the difficulty of the three, six, nine and twelve line sets should be equal. The cards chosen are shown in Table 3.

For the experiment on group influence, it had to be decided which line on each card was to be chosen as the 'group response': it was decided to choose the most frequently given incorrect line (by the twenty subjects in the pre-test experiment) as the group response, and to count as conformity an increased number of responses of these lines by critical (i.e. 'group-influence') subjects as compared with subjects in the display design pre-test.

The suggested definition of ambiguity is in terms of the probability of a stable perception resulting from a stimulus configuration. It is not clear how this should be measured objectively and a tentative index of ambiguity was calculated by subtracting the probability of the line chosen as the group response (as described above) being seen as correct in the individual situation from the highest probability of any particular correct line being seen as correct in the individual situation. The lower this difference the higher the ambiguity.

On Table 3 these differences in probability are shown. It can be seen that in the  $\frac{1}{16}$  and  $\frac{1}{4}$  in. difference cards the ambiguity is not reliably higher for the six-, nine- and twelve-line cards than for the three-line cards. The hypothesis has to be amended, therefore, for these sets of cards in that it would not be predicted that higher conformity would occur in the multi-answer cards. It would be predicted in the  $\frac{3}{8}$ ,  $\frac{1}{2}$  and  $\frac{3}{4}$  in. difference cards but not significantly differently as between six-, nine- and twelve-line cards in the  $\frac{3}{8}$  and  $\frac{1}{2}$  in. differences, since the ambiguities do not differ significantly.

Table 3. *Card numbers and differences (in parentheses) between highest correct response probability and 'group response' probability*

Size difference (in.)	No. of lines			
	3	6	9	12
$\frac{1}{16}$	21 (-0.2)	2 (-0.2)	3 (-0.1)	24 (-0.1)
$\frac{1}{8}$	5 (0)	26 (-0.2)	7 (0)	8 (0)
$\frac{3}{8}$	9 (0.3)	30 (0.15)	11 (0.2)	32 (0.1)
$\frac{1}{2}$	13 (0.4)	14 (0.15)	35 (0.15)	16 (0.15)
$\frac{3}{4}$	37 (0.9)	18 (0.5)	39 (0.25)	40 (0.25)

For the group experiment it was thought best not to have an unbroken sequence of twenty unanimous incorrect responses to all twenty cards, so five more cards were added, three at the beginning of the series (to the first and third of which the group responses were unanimous and correct, and to the second a divergent response was given by one of the group), and two others later in the series to which divergent responses were given by one member of the group. Only the cards to which the group gave a 'unanimous' incorrect response were considered in calculating the relative conformity values of the 'critical' subjects' responses.

#### *Group experiment*

It was decided to use influencing, or 'environment', groups of five subjects, making, along with the critical subjects, total groups of six.

The influencing group had a training period before the experiment proper began; they were interested in the experiment and were adequate actors. They were warned not to arrive in company at the experimental session, and not to betray familiarity with the experimental arrangements.

When the influencing group had been trained, twenty 'critical' subjects were selected and assigned to various experimental sessions. They were told that an experiment in judging was being carried out which required large numbers of subjects and, therefore, the subjects would be tested in groups. In each case they were told that a group was nearly complete and they would be assigned to this group and so complete it.

The physical situation was the same as in the individual experiments. Thus the subjects sat at a large table at the other side of which the stimulus frame was fixed. Six chairs were set out on one side of the table and in front of each chair, on the table, a foolscap sheet was laid with the subject's name and the experiment number written on. These named sheets were set out so that the critical subject sat fifth from the left.

The subjects came into the laboratory and were asked to sit at the place where they found their name. Behind them an experimenter sat at a table and another experimenter operated the stimulus frame.

The instructions given were:

'I am going to show you two cards, one of which has one line on it and the other several lines. Your task is to decide which of the several lines is equal in length to the single line on the other card. The lines have letters printed under them and the response you give is the letter under the line you think equal. If you give your responses in turn from left to right, they can be recorded by the experimenter sitting behind you. If you turn over the sheets in front of you, you will find a schedule for recording your responses yourself. We would be glad if you would do this as it provides a check on our record. When you have all responded I will remove the two cards and replace them with two others. The standard line is of equal length in all cases.'

It was necessary to have the subjects record their responses because this was the only way in which it was possible to give the first group subject (the one on the extreme left) a list of the responses he was to make. They were pencilled faintly on his response sheet. All the critical subjects accepted this double response situation as realistic according to their later subjective reports. The first two stimulus cards were then presented, the first environment group subject gave his response, the second gave the same response, and so on. When they had all responded the next stimulus was presented, and so through the set of twenty-five cards.

When the experimental series was complete the environment group subjects left the laboratory and the critical subject was asked to remain. The purpose and nature of the experiment was then explained and a subjective report obtained.

### Quantitative results

On Table 4 can be seen the number of responses which agreed with the group response on each card and on Table 5 the differences between the number of these responses and the number given in the individual situation. A 't' test comparing the mean of these differences with zero is highly significant (Table 5).

Table 4. Card numbers and conforming responses (in parentheses) by twenty critical subjects which agreed with the 'group' response

Size difference (in.)	Number of lines			
	3	6	9	12
$\frac{1}{16}$	21 (11)	2 (10)	3 (6)	24 (9)
$\frac{1}{8}$	5 (13)	20 (14)	7 (8)	8 (5)
$\frac{1}{4}$	9 (9)	30 (5)	11 (6)	32 (6)
$\frac{3}{8}$	13 (11)	14 (4)	35 (6)	16 (6)
$\frac{1}{2}$	37 (6)	18 (5)	39 (6)	40 (1)

Therefore, this experiment can be assumed to have produced conformity, as a result of group influence.

But an examination of Table 5 shows that the hypothesis was not supported in that there is no significantly greater conformity on the multi-answer cards (in fact slightly less) and there is no greater conformity on the smaller size difference cards. In fact there is no relationship between the possible ambiguity values of Table 3 and the incidence of conformity responses.

Table 5. Differences between incidence of conforming responses and frequency of same response in individual situation

Size difference (in.)	Number of lines				Total
	3	6	9	12	
$\frac{1}{16}$	2	3	0	4	9
$\frac{1}{8}$	3	5	5	1	14
$\frac{1}{4}$	2	0	3	2	7
$\frac{3}{8}$	5	-2	3	4	10
$\frac{1}{2}$	5	4	4	0	13
$\frac{3}{4}$	5	4	15	11	
Total	17	10	15	11	

$t_{19} = 6.16$ , probability of this occurring by chance =  $1 \times 10^{-2}$ .



*Subjective reports and observations*

All subjects accepted the situation as genuine. Three had heard about social influence on behaviour from some source or other and one of these considered the possibility that the experiment was concerned with this but decided against it.

Those subjects who had generally not conformed were relieved to know that the situation had been a contrived one because it meant that the evidence of their perceptions was still reliable. The conforming subjects, naturally, had more mixed feelings and their general reactions were that the task was very difficult and that they thought they might as well agree. When it was pointed out that individuals were able to make correct responses to about half the stimuli, on average, they usually said that they suffered from what seemed to be perceptual confusion. This is similar to some of the results obtained by Asch. One subject who gave the same response as the group to eighteen stimuli claimed that he had not conformed at all! So far as one could judge the perceptual confusion resulting from the stimuli was such that he genuinely felt this to be true. On the other hand, his behaviour during the experiment appeared to indicate anxiety, and it is difficult to assess the status of his conformity.

In general, the behaviour of the critical subjects during the experiment was related to the way they dealt with the situation in terms of responses. It is not possible to be quantitative about this data but the general impression was that those subjects who tended to conform showed anxiety, subdued behaviour, and what looked like a desire to efface themselves in the group; while those who resisted group influence removed themselves from the group physically as far as was possible (e.g. tilted their chairs back or moved them back out of line), gave their responses in a loud voice, and looked round in a bored manner between stimuli.

The subjects were more or less all conscious of the possibility of being influenced, but were often not conscious of having been influenced. They were, in general, disturbed by the difference of the unanimous group response from the response they gave or intended to give.

One subject had had a lot of experience of judging the relative lengths of lines, as it had been part of his job at some time to examine the amplitude of wave-form pen-recordings. This subject made only one conforming response and showed amusement at the responses of the group.

The third subject tested said in the course of the interview that the group were more likely to be correct in their judgement, in his opinion, when the stimulus card had only three lines than when it had more than three lines. The remaining seventeen subjects were asked whether in their opinion the group seemed more likely to be correct when there were three lines or more than three lines. Fifteen subjects said when there were only three lines, one said when there were more than three lines and one said it did not make any difference. Counting the last two as 'more than three' and assuming equi-probability for 'three' and 'more than three' the chances of obtaining fifteen out of seventeen responses of 'three' is approximately 0.001. Therefore, it is fairly safe to reject the assumption of equi-probability and to assume that this represents a genuine opinion of the subjects.

Some of the subjects reported that they felt 'stubborn' when they gave a response

different from the group. Thus even though the group had been instructed to behave as neutrally as possible (and on the whole they achieved this) the critical subjects tended to project disapproval onto the group.

### *Consideration of results*

As can be seen from Table 5 the hypothesis was not substantiated by this experiment, as there were no more conforming responses on the six-, nine-, and twelve-line stimuli than on the three-line stimuli.

The reason seems to be that the experimental procedure was one which did not adequately test the hypothesis.

The subjects questioned on the point (seventeen out of twenty) thought that the group were more likely to be correct when the stimulus had only three lines than when it had more than three lines. Logically, of course, this means that the critical subjects themselves were more likely to be correct in this situation, but it does not seem to have been seen in this way. As far as one can judge, the subjects had an opinion about the line to choose and the group had an opinion; when the group was more likely to be right the critical subjects were inclined to agree with the group regardless of the status of their own estimate. The situation need not have been of this kind, of course; their own estimate could have subjectively increased in status as the number of lines decreased. But it seems that the group's influence increases with increasing subjective ease of task, at a different rate to their own certainty. Of course, the critical subjects were correct in their assumption that when there are only three lines, one of which is correct, the group was more likely to be right than when there were more, if there was only one correct line. They were not aware either that the group was responding to instruction or that there was more than one correct answer in the six-and-more-lines stimuli.

This clearly alters the nature of the experiment. The assumption was that the group influence would act with equal force for all the stimuli, or that the differences would be random. A systematic difference of this kind means that the three-and-more line cards were not equal in the amount of influence which was felt, and this difference in influence operated directly against the possible operating of the hypothesis, i.e. this difference in influence favoured conformity on the three-line cards and the hypothesis effect would favour conformity on the more-than-three-line cards.

It will be noticed from Table 5 that conformity did not increase as the size difference between the lines decreased. As the size difference decreased the task became objectively more difficult (Table 2) and hence more ambiguous in the usual sense. It was predicted that conformity would increase as the size differences decreased since the task was then becoming more difficult. During the interviews with the critical subjects after the experimental sessions they were not questioned about where they felt the group's influence most in terms of size difference, but it is tempting to assume that where the size differences could be seen to be large the assumption was made that the group was more likely to be right than where they could be seen to be very small.

As can be seen from Table 5 there is no significant difference in the incidence of conformity responses either in the different size differences or the different numbers of lines.

It seems plausible to consider both these stimulus dimensions (size difference and number of lines) as dimensions in which increasing ambiguity causes an increase in conformity but at the same time decreasing subjective weight given to the group response causes a decrease in conformity.

It is not claimed that this constitutes sufficient evidence to support the hypothesis, but since it is known that the group's opinion seemed more likely to be right when there were only three lines and there is not a significantly greater conformity on the three-line cards, and we know that difficulty increased as the size differences decreased but there is no significant increase in conformity as the size differences decrease, this is one possible interpretation.

We know sufficient, however, for it to be clear that the status of the group's response varies with the quality of the stimulus as seen by the critical subject. Deutsch & Gerard (1955) distinguish between what they call 'normative social influence' which is the influence to conform to the positive expectations of another and 'informational social influence' which is the influence to accept information obtained from another as information about reality.

Presumably both these influences operated in the present experiment and it is clear that the status of the 'informational social influence' varied considerably with the nature of the stimulus. The group's responses were known to be a function of the stimulus situation and where the stimulus situation was such that it was difficult to get information then the group's responses had less 'informational social influence'. It is not clear whether normative social influence varied in the same way, although it is possible that where the group was more likely to be right, and hence could be presumed by the critical subject to feel more certain of their judgement, then their expectation of agreement could also be presumed to be higher, and so the normative social influence they exerted would be greater.

This means that this experiment, and conformity experiments of this general kind, are not equivalent to prestige suggestion situations.

Prestige presumably has two components, (a) a person having prestige is a person whose expectations of agreement a critical subject would not care to disappoint—that is a person with prestige exerts high normative influence, and (b) a person with prestige has superior information or can easily be presumed to have superior information about the real nature of the stimulus situation (he exerts high informational social influence).

In the present experiment all the subjects apparently came into the situation together, so all had the same experience and the stimulus was simply a set of lines whose lengths had to be judged—a familiar enough situation—so the informational social influence was quite weak. Its main strength resulted from the fact that the group were unanimous (the combined estimate was perhaps better than any single estimate), but not from any superior experience or capacity of the group.

This, of course, did not mean that the informational social influence did not exist. As Deutsch & Gerard point out, the normal experience of human beings is that others' perceptions and their reports of these are generally reliable, i.e. agree with the veridical situation and with the experience of the subject making the assessment, and it is not reasonable to expect that the stimulus will be perceived in exactly the same way when the description of other people does not agree with one's own experience.



The importance of the relative efficiency of the critical subjects and environment group subjects in their carrying out of a task has been shown by Samelson (1957) to have an important bearing on the incidence of conformity. He used two groups of critical subjects in tachistoscopic recognition tests. One group of critical subjects had been given the implicit information that they were more efficient at the general task of tachistoscopic recognition than their environment groups, and these critical subjects conformed significantly less than the other group of critical subjects which had not been given this information. Di Vesta (1959) found that the degree of conformity decreased as the frequency of errors made by the majority increased.

So in general it is clear that informational social influence is a function of the probability that the group is correct and in particular is a function of the ambiguity or difficulty of the stimulus, so that the subject's uncertainty and hence tendency to be influenced is compensated for by his doubt of the group's correctness.

### *Subjective certainty*

The subjective certainty of the subjects about the correctness of their own response is an important variable. It may be that the certainty of the subject is inversely correlated with ambiguity and is in fact the psychological variable equivalent to variations in ambiguity.

Since the certainty of the subjects in the group experiment varied systematically with the nature of the stimulus material it was decided to carry out a further test of the effects of the material on subjects in the individual situation.

The experiment carried out was of essentially the same form as the previous display design experiment. The stimuli were presented to sixteen individual subjects on two occasions separated by about 2 weeks. On one occasion an estimate of their certainty was obtained by asking the subjects to score each response they gave on a five-point scale of certainty ranging from absolutely certain to very uncertain.

On the other occasion the response time for each stimulus was noted by asking the subjects to press a key when they had decided on their response. This stopped an electronic timing device which was started by the display being put into position. This technique, of course, does not give reliable absolute response times, but it was felt that it would be a sufficiently accurate procedure to give an indication of the relationship between response times to different stimuli.

Half of the subjects had the certainty experiment first and half the response time experiment first. It was predicted that certainty would decrease as the number of lines increased and as the size differences decreased. Also that response time would increase as the size difference decreased and the number of lines increased. The results in terms of average certainty and average response time for each card are shown in Tables 6 and 7.

The predictions were tested by noting the difference between successive cards in the set. There were for the 'certainty-no. of lines' prediction fifteen differences between successive cards, i.e. between card numbers 1 and 2, between card numbers 2 and 3, and so on to cards 39 and 40. Of these differences, every one was in the predicted direction. For 'certainty-size differences' the score was 13 out of 16 as predicted, for 'response times-number of lines' 15 out of 15 and for 'response times-size differences' 13 out of 16 (Table 8).

Table 6. Card number and average certainty (in parentheses)

Size difference (in.)	No. of lines			
	3	6	9	12
$\frac{1}{16}$	21 (2.27)	2 (2.00)	3 (1.75)	24 (1.64)
$\frac{1}{4}$	5 (2.89)	26 (2.48)	7 (2.18)	8 (1.95)
$\frac{3}{8}$	9 (3.31)	30 (2.70)	11 (2.37)	32 (2.23)
$\frac{1}{2}$	13 (3.40)	14 (2.70)	35 (2.31)	16 (2.18)
$\frac{3}{4}$	37 (3.73)	18 (3.00)	39 (2.52)	40 (2.37)

Table 7. Card number and average response time (in parentheses) in seconds

Size difference (in.)	No. of lines			
	3	6	9	12
$\frac{1}{16}$	21 (7.64)	2 (10.32)	3 (11.15)	24 (13.51)
$\frac{1}{4}$	5 (5.69)	26 (9.19)	7 (10.01)	8 (14.52)
$\frac{3}{8}$	9 (4.80)	30 (7.44)	11 (10.63)	32 (12.83)
$\frac{1}{2}$	13 (3.75)	14 (7.57)	35 (9.22)	16 (11.20)
$\frac{3}{4}$	37 (2.57)	18 (6.29)	39 (9.19)	40 (10.91)

Table 8

Certainty prediction	No. of differences	No. as predicted
(a) Will decrease as no. of lines increases	15	15
(b) Will increase as size difference increases	16	13
Response time predictions		
(a) Will increase as no. of lines increases	15	15
(b) Will decrease as size differences increase	16	13

This confirms that certainty is inversely related to ambiguity as defined, and response time positively related.

The interesting thing, to go back to the 'group-influence' experiment, is that decreasing certainty (increasing ambiguity) does not lead to increasing conformity because the informational social influence exerted by the group is seen to be related to the characteristics of the stimulus.

But the critical subjects generally tended to use their opinion of the nature of the stimulus and other factors in the situation to assess the status of the group response rather than the status of their own opinion. For example, it was noticed (although there is no quantitative analysis of this) that where the group gave their responses quickly, which would presumably suggest that they were certain of the correctness of their response, the critical subject tended to agree more than where the group's responses followed one another more slowly.

*Effects produced by social influence*

A question which is of central importance is the precise nature of the psychological change brought about in the critical subject by social influence. In the suggestion experiments it appears to be accepted that the change is in the nature of the subjects' perception; the subject sees (as a result of suggestion) something different from what he would have seen without it. In the Sherif (1935) experiments, on the other hand, the general impression is that a judgemental change took place, the explanation

being that a change in the 'frame of reference' had occurred, and, therefore, the stimulus was judged differently.

If the judgement of a stimulus changes as a result of social influence then either the frame of reference or memory of past stimuli must have changed, or the immediate perception of the stimulus must have changed.

It seems that social influence would be much more fundamental in its nature if it changed, for example, the present memory of total past experience of length, than if it changed the immediate perception of a particular stimulus or series of stimuli.

This may be a mistaken analysis in one respect in that the change which occurs may be one of naming only. The frame of reference may be stored in some form which makes it possible for a subject under social influence to change the names he gives to different remembered lengths without in any way changing either the distribution of remembered lengths or the perception of the present stimulus.

It is not likely that the problem of the exact psychological process resulting in change of response will be easily solved, but it is suggested that the change is in the nature of a local distortion in an area of experience rather than in a large-scale shift of anything that could be called a frame of reference.

For example, although the Sherif situation is further complicated because the perception is in any case an illusion—i.e. not only is it impossible to study the subjects' perceptions or judgements directly, it is impossible to study the stimulus because there was in fact no movement—it might be that the information from other subjects caused the perception of all the movements to change in that they were seen to be at a different distance (an equal angular movement at a greater distance would be a greater linear movement and vice versa). This would not be equivalent to a long-term change of frame of reference but to a change in one aspect of the nature of the immediate stimulus which would, however, be common to all the stimuli.

It was hoped in the present experiment to obtain some evidence on this matter, perhaps by getting different types of subjective reports related to different incidences of conformity. Nothing very significant of this kind emerges, however, except that as far as it is possible to judge perceptual confusion was produced in some of the subjects.

Table 9. *Distribution of responses by size of line excluding conforming responses of critical subjects and equivalent responses by individual subjects*

Cards with standard at bottom			
	Individual subjects		Critical subjects
	%		%
	107 85		87 82
Standard			
Length equal to group response	10	7.9	12 11.3
Standard +	9	7.1	7 6.6
Standard + +	126		106
Cards with standard at centre			
	18 10		9 6.1
Standard -	137 76.1		112 76.1
Standard			
Length equal to group response	25	13.9	26 17.7
Standard +	180		147



There is some slight quantitative evidence, however. In Table 9 certain responses of the individual subjects (those used in the display design experiment) and the critical subjects are compared. The responses in question are those which both groups made which are not equal to the influencing group response. Of course, the individual subjects were not exposed to an influencing group but those responses they made which were the line later chosen as the influencing group response are not included.

Table 9 is in two sections; the first section is concerned with the nine cards on which the line equal to the standard was at the bottom of the distribution of three lengths on the stimulus cards, and the second section with those cards on which it was at the centre of the distribution.

The responses are categorized according to whether they were equal to the standard, above it, below it, or two stages above.

It can be seen that in the first section there is a slight increase in the incidence of responses of the critical subjects compared with the individual subjects of lines one stage above the standard. These lines were those equal in length to the group response. In the second section the same effect can be seen.

The differences are not significant and are only worth noting because the same effect occurs on both types of cards although the process involved is different; in the first section there is an increase in responses of the central line and in the second section an increase in responses of the largest line.

The only common factor is that both these types of lines are of the same length as the line the group gave as their response.

The conforming responses are not included, of course, so this slight increase represents a tendency of the critical subjects to give as their response a line equal in length to the group response, even when they did not agree with the group response.

It is suggested that this represents a perceptual distortion in the critical subjects in so far as the relative lengths of standard and stimulus lines were affected by the group responses in such a way that critical subjects tended to give a larger line as 'equal' to the standard, even when continuing to respond on the basis of what they saw, rather than merely agree with the group response.

This paper is based on an M.Sc. thesis submitted to Edinburgh University and the research was supported by D.S.I.R.

#### REFERENCES

- ASCH, S. E. (1956). Studies of independence and conformity: 1. A minority of one against a unanimous majority. *Psychol. Monogr.* **70**, no. 416, 3-9.
- COFFIN, T. E. (1941). Some conditions of suggestion and suggestibility: A study of certain attitudinal and situational factors influencing the process of suggestion. *Psychol. Monogr.* **53**, no. 241, 81-116.
- DEUTSCH, N. & GERARD, H. B. (1955). A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.* **51**, 629-36.
- DI VESTA, F. J. (1959). Effects of confidence and motivation on susceptibility to informational social influence. *J. Abnorm. Soc. Psychol.* **59**, 204-209.
- SAMELSON, F. (1957). Conforming behaviour under two conditions of conflict in the cognitive field. *J. Abnorm. Soc. Psychol.* **55**, 181-7.
- SHERIF, M. (1935). A study of some social factors in perception. *Arch. Psychol.* no. 187.

(Manuscript received 26 July 1961)

## CONDITIONING AND PERSONALITY

By H. J. EYSENCK

*University of London*

The application of learning theory to the study of personality presents many problems, some of which are here discussed in reply to a critique by Champion of the author's particular contribution to this field. It is suggested that such discussions are relatively fruitless unless they take into account the differences between 'strong' and 'weak' theories, and unless they bear in mind the many different ways along which predictions may be mediated from postulates.

Champion's (1961) recent comment on the writer's use of learning theory in relation to personality makes a number of points which could all be argued at length. Similarly, there have been other critics who have been concerned with other points of the general theory put forward in *Dynamics of Anxiety and Hysteria* (Eysenck, 1957). Some of these criticisms arise from a certain confusion which is almost inevitable when certain notions originally advocated by Hull have to be used with an alteration in meaning and content made necessary by more recent experimental findings. 'Reactive inhibition' in the writer's theoretical framework is clearly not peripheral and work-produced as it is for Hull, but central and not crucially related to the actual amount of physical work done. Similarly, Hull makes  $I_R$ , conceived as a negative drive additive with  $sI_R$ , a habit; the writer has followed Gwynne Jones (1958) in subtracting  $I_R$  from  $D$  instead. The term 'excitation' has been used in the Pavlovian sense, i.e. as a facilitating factor in neural transmission, perception, and learning rather than in Hull's sense of  $sE_R$ . Failure to take account of these and other changes in meaning make some of the criticisms levelled at *Dynamics of Anxiety and Hysteria* inapplicable.

A more important source of misunderstanding may be the conception of the role of theory in science held by the writer (Eysenck, 1960a). Hull, Champion and many other learning theorists appear to regard learning theory as what has been called a 'strong theory', capable of mediating precise quantitative deductions along rigorous lines. While this is no doubt a most desirable type of theory to have, it seems to the writer that at present the only theories available to psychologists are 'weak' theories which at best suggest areas of exploration, possible lines of advancement, and the outlines of experimentally testable nomological networks (Eysenck, 1960a). Weak theories of this type cannot reasonably be criticized along lines which would be appropriate for strong theories, and confusion between the two may account for a good deal of the rather futile warfare between those who label themselves as 'pro' theory and those who regard themselves as 'anti' theory.

The main advantages of the weak theory are (1) that it may lead to the discovery of new and unexplored facts, and (2) that it may raise new problems which had not previously been considered at all. In due course weak theories may turn into strong ones by pursuing the paths suggested by these new discoveries, and by the carrying out of experiments along these new lines; in this process there will inevitably occur a transmutation which will leave very little of the original theory intact, except

perhaps as a way of looking at the field, and in the way of certain general concepts. In the theory put forward in *Dynamics of Anxiety and Hysteria* the concepts of inhibition and excitation were linked in certain ways with extraversion and introversion, and used to make certain predictions, some of which were later verified experimentally; it was not the writer's intention that inhibition and excitation should be regarded as concepts having a perfectly rigorous definition and meaning, but it was his hope that the relations indicated by the theory would help in clarifying the nature and meaning of these concepts (Eysenck, 1962).

It may be possible to exemplify the way in which the writer has attempted to use learning theory by taking as a particular example the relationship between conditioning and personality. Starting with a theory about the existence of two important personality dimensions, neuroticism or emotionality and extraversion-introversion (Eysenck, 1960*b*), we went on to seek for causal factors to account for a given individual's position on these dimensions. Emotionality may be identified, with some misgivings, as a consequence of an over-labile autonomic system; extraversion was conceived of as being related to the notions of inhibition and excitation as used and operationally defined by Pavlov and Hull. It seemed that in order to mediate predictions from the conceptual and experimental levels to the behavioural level it was necessary to postulate some such mechanism as 'defective conditionability' in extraverts, and 'enhanced conditionability' in introverts; as will be shown later there are many different ways in which predictions of this kind could be derived from the hypothesis linking extraversion-introversion and the excitation/inhibition balance.

There was available at the time no single experimental report of any such connexion, although there were several experiments purporting to verify the Spence-Taylor hypothesis relating conditionability and scores on the Manifest Anxiety Scale. This scale is a good measure of neuroticism, correlating very highly with such scales as the MPI Neuroticism scale, or the Maudsley Medical Questionnaire; it also correlates positively, although much less highly, with introversion (Eysenck, 1957). When correction is made for attenuation due to unreliability of the scales, it can be shown that scores on the MAS can be predicted fairly exactly from scores on the MPI, the Neuroticism scale contributing some 80% and the Extraversion scale (reversed) contributing less than 20%. The Spence-Taylor notion of anxiety as a drive seemed to identify their predictions of higher conditionability with the dimension of neuroticism; the writer's theory would account for their findings in terms of the (small) introversion content of the MAS. We thus have two clearly different predictions, relating 'conditionability' respectively to *neuroticism* or to *introversion*; work done on the MAS is largely irrelevant in this connexion as this scale partakes of both these orthogonal dimensions. (It would of course be relevant if the findings had been consistently negative, as this would have contraindicated both theories.)

At the writer's suggestion, the crucial experiment was carried out by Franks (1956, 1957) who used both normal and neurotic introverts and extraverts. Taking both groups together, he obtained results which show that in the eighteen test trials interspersed among the conditioning trials introverts show a proportion of conditioned responses over twice as large as that shown by extraverts; the scores of 35 subjects in each group are given in Fig. 1. Similar results have been obtained with the GSR



by Vogel (1960); she found that introverts required a mean of 5.18 trials to conditioning, while extraverts required 12.25 trials. Franks failed to find any relationship between conditioning and neuroticism in either of his groups; Vogel found a barely significant one. Halberstam (1961), using PGR conditioning on normal controls, hysterics and psychasthenics, found that the introverted neurotics conditioned in half the number of trials (19.61) needed by the extraverted neurotics (40.94); the normal controls were intermediate (23.33). These findings 'fully agree with the similar results obtained by Franks'. Hysterics were also found to extinguish more quickly, both under 'informed' and 'uninformed' conditions. (Cf. also Barendregt, 1961,

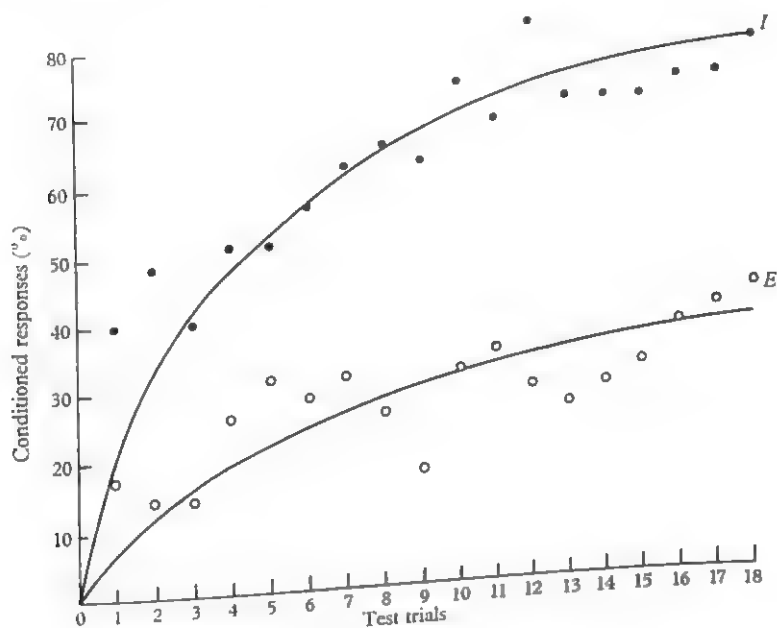


Fig. 1

p. 211, who found a correlation of 0.29 between eye-blink conditioning and introversion.) These findings seem to indicate that the predicted relation between conditioning and introversion is not entirely absent, and that if these results can be duplicated by other research workers, then our theory has passed the main test of a scientific theory, to wit, *the mediation of prediction leading to new and previously unexpected facts*. It also fulfils a second test, that of accounting for the known facts, such as the correlation between conditioning and the MAS. It is not necessary at this point to discuss the now voluminous literature bearing on this point; most studies give results in the predicted direction, but these are not always significant. Thus Das (1957) failed to find a significant correlation between introversion and conditioning, but his subjects were partly white, partly coloured and non-European; inspection of the questionnaire responses of the two groups showed that they could on no account be regarded as coming from the same universe. Field (1960) obtained rather lower correlations between introversion and conditioning than had Franks, but he was working with a prison population to whom the ordinary MPI with its stress on sociability questions does not apply particularly well.

When it comes to the precise mediation of our hypothesis in terms of learning theory, a large number of possibilities must be considered, and it is quite impossible to make a rational choice at present. Consider these possibilities. (1) Introverts have greater excitatory potential, facilitating  $sH_R$  acquisition. This would directly lead to the correct prediction regarding the higher conditionability of introverts; some evidence in favour of this hypothesis is given by Eysenck (1960c). (2) Extraverts have greater inhibitory potentials ( $I_R$ ). This could be made to lead to the correct prediction of higher conditionability of introverts along several different lines: (a)  $I_R$ , being a negative drive, subtracts from  $D$  according to the Gwynne Jones formula (Gwynne Jones, 1958), leading to lower effective drive  $\bar{D}$ . This deduction would bring drive into the picture, but along quite a different route from that favoured by Spence and Taylor. (b)  $I_R$ , attaching to the stimuli used in the experiment, leads to habituation and consequently to a lowering of the *effective* strength of the UCS. (c)  $I_R$ , generalizing over the cortex, leads to 'sleep-inhibition', or loss of 'arousal'; there is some direct evidence for both (b) and (c) in a recent paper by Voronin, Sokolov & Bao-Khua (1959). (d) The autonomic responses associated with the stimuli become habituated due to  $I_R$  attaching to them and thus lower the total effective drive-arousal combination. This deduction also would lend to the prediction of a lessening of drive, again quite different in origin to the Spence-Taylor hypothesis. Certain subjective observations made by Franks (1956) and by other investigators in our laboratories indicate that one or all of these possibilities may be true; thus Franks remarks that 'the poor conditionability of the hysterics, their more rapid PGR adaptation to the air puffs, their subjective reports that the air puffs were not very disturbing, and perhaps their reports of feeling sleepy all support the hypothesis that hysterics are in a state of cortical inhibition'.

It should not be assumed that these possibilities exhaust the supply; many others will occur to the reader. It will be a long time before the precise chain of causation is known, and much precise and detailed work will be required. It is unlikely, as Champion apparently believes, that results such as those plotted in Fig. 1 can be used neatly to separate our rate of learning and drive strength. If the reader will consult Champion's Figs. 1, 2 and 3, which illustrate the consequences on conditioning of variations in rate of learning, in drive strength, and in both jointly, he will notice that a decision in any particular case only becomes possible (if then) when the curves are known to have reached their respective asymptotes. If now the reader will draw a line parallel to the ordinate intersecting the abscissa about  $\frac{1}{2}$  in. from the origin, and consider the curves up to this point, he will find that they all look pretty much alike, and indeed that they all look rather like those given in Fig. 1. In other words, where the curves clearly fall short of the asymptote, no decision between these various explanations is possible. Eysenck (1957) did provide a discussion and some tentative mathematical manipulations of the factual observations, but as clearly pointed out on p. 119, these depend on extrapolations which of necessity are very uncertain indeed. Champion has taken these *jeux d'esprit* too seriously; they were not intended to prove anything, but merely to indicate lines along which research might profitably proceed.

While thus there are many possibilities of interpretation of the observed results, it would seem that the original Spence-Taylor hypothesis is not strongly supported.

The evidence suggests a definite relationship between conditioning and introversion, little or none between conditioning and neuroticism. (More recent work by Becker & Matteson (1961) may require a reappraisal of this conclusion.) Champion says that the members of the Iowa group 'have not, as Eysenck suggests, seriously attempted to relate learning theory to personality functioning'; but it has never been maintained that they had. Quite the contrary; Eysenck has criticized the Spence-Taylor group for failing to consider the well-substantiated knowledge regarding personality structure now available (Eysenck, 1957), and thus arbitrarily settling on a questionnaire (the MAS) without any knowledge of its dimensional structure. They thus failed to perform what must be regarded as the crucial experiment in the argument between them and the London group, to wit, the examination of subjects *high* on neuroticism (and the MAS) and *low* on introversion. The hysterics in Franks's group fulfilled this qualification, and as will be remembered, showed low conditioning in spite of their high MAS scores. Similar findings are reported by Lykken (1957), working with psychopaths and sociopaths, as well as by Tong & Murphy (1960), who refer to the low conditionability of psychopaths as 'an accepted laboratory fact' (p. 1285). The evidence against the Iowa group, and in favour of the London group, seems sufficient to indicate that the suggestion of a significant relation between introversion and conditioning may not be altogether mistaken.

To say this does not, of course, rule out the possibility that under certain circumstances neuroticism too may be found to be related to conditionability, or that introversion may fail to be so related. It would be naïve to believe (although it is often useful and time-saving to write and talk as if this were so) that there is a specific and permanently demonstrable relation between 'conditioning' and a given personality trait. The term 'conditioning' is not closely enough defined to carry any predictive burden, and any proposition stated so broadly cannot in the nature of the case be disproved or supported. (1) In the first place, different types of conditioning (eye-blink GSR, heart rate, hand-withdrawal, skin temperature) do not usually intercorrelate very highly, suggesting the existence of considerable response specificity (perhaps analogous to Lacey's (1950) concept of autonomic response specificity.) (2) Within a given type of conditioning, there are many important parameters which must be explored in detail in order to discover their relevance to the hypothesized relationship. Thus in eye-blink conditioning the CS-UCS interval is usually taken at the population optimum of about 450 msec.; it is quite in line with theory to assume that introverts and extraverts have different optima, so that variations in this interval from experiment to experiment may change the nature of the relationship between conditioning and personality. (Some indirect evidence for this view is offered by Eysenck (1962).) (3) Both Franks and the Spence-Taylor group use a fixed strength of puff to the eye as the UCS. But individual differences should be taken respect to their lid-closure thresholds, and such individual differences should be taken into account; our more recent studies along these lines support the postulation of considerable individual differences in this threshold. (4) Experiments may differ in the degree of *awareness* the subject is allowed of the purpose and the nature of the experiment. We have recently compared the traditional type of eye-blink experiment with one in which subjects were instructed to respond with a depression of a key to the CS; UCS was introduced as a punishment for slow reactions. Awareness is thus



reduced, and altogether higher rates of conditioning achieved (Issa, unpublished data). (5) Experiments differ in the degree of *spatial inhibition* allowed (distraction, etc.). If extraverts, as postulated, show greater degrees of spatial inhibition, then manipulation of this variable may be important in accounting for different experimental results. (6) The strength of the UCS differs from experiment to experiment; it is not impossible that with *low* values introversion shows higher correlations, with *high* values neuroticism. (7) Experiments differ in length of total trial, length of inter-trial pauses, spacing of such pauses, as well as in the use of separate test trials, interspersed with training trials. It may be surmised that experimental arrangements will favour introverts more the more massed the trials are, and the greater is the number of unreinforced test trials, as compared with reinforced training trials. These are only some of the parameters which may distinguish one 'conditioning' experiment from another, and it is by no means reasonable to assume that correlations between conditioning and personality achieved with one combination will be duplicated with another. Clearly, the search for a satisfactory conception of 'conditionability' and its relation to personality variables is only at the beginning; it may be surmised that the outcome will throw important new light, not only on personality development and breakdown, but also on the very concept of conditioning itself.

In conclusion, it seems worth while to draw attention to Champion's reticence in refusing to deal with 'the fairly extensive sections of Eysenck's book which relate to inhibition, work decrement, reminiscence, and similar factors'. The theory put forward derives whatever interest it may have from the fact that it tried to relate a large number of experimental facts to a limited set of explanatory variables, and these in turn to personality and drug action; in each particular case many different alternative explanations of the experimental findings are possible. It is when the total set of related data is viewed as a whole (which as we all have learned long ago is more than the sum of the parts) that such alternative hypotheses, which never cover more than one, or at most a few of the experiments, cease to be attractive. One's judgement of the possibility of explaining the lesser conditionability of extraverts in terms of cortical inhibition is probably increased by the fact that such inhibition can be demonstrated in many other circumstances, and in connexion with many other types of experiment. To isolate one phenomenon and treat the rest of the evidence as non-existent is not likely to give the reader the correct perspective for judging the theory under discussion; '*natura in reticulum sua genera connexit, non in catenam*' (Haller, 1768).

#### REFERENCES

- BARENDREGT, J. P. (1961). *Research in Psychodiagnostics*. The Hague: Mouton.  
 BECKER, W. C. & MATTESON, H. M. (1961). GSR conditioning, anxiety and extraversion. *J. Abnorm. Soc. Psychol.* **62**, 427-30.  
 CHAMPION, R. A. (1961). Some comments on Eysenck's treatment of modern learning theory. *Brit. J. Psychol.* **52**, 167-73.  
 DAS, J. G. (1957). An experimental study of the relation between hypnosis, conditioning and reactive inhibition. Unpublished Ph.D. Thesis, University of London.  
 EYSENCK, H. J. (1957). *Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.  
 EYSENCK, H. J. (1960a). *Experiments in Personality*. London: Routledge and Kegan Paul.  
 EYSENCK, H. J. (1960b). *Structure of human personality*. London: Methuen.

- EYSENCK, H. J. (1960c). Objective psychological tests and the assessment of drug effects. *Int. Rev. Neurobiol.* **2**, 333-84.
- EYSENCK, H. J. (1962). *Experiments with Drugs*. London: Pergamon Press.
- FIELD, J. G. (1960). Eyelid conditioning and three personality parameters. Paper read at the XVI Int. Congr. Psychol., Bonn.
- FIELD, J. G. & BRENGELMANN, J. C. (1961). Eyelid conditioning and three personality parameters. *J. Abnorm. Soc. Psychol.* **63**, 517-23.
- FRANKS, C. (1956). Conditioning and personality: a study of normal and neurotic subjects. *J. Abnorm. Soc. Psychol.* **52**, 143-50.
- FRANKS, C. (1957). Personality factors and the rate of conditioning. *Brit. J. Psychol.* **48**, 119-26.
- HALBERSTAM, J. L. (1961). Some personality correlates of conditioning, generalization, and extinction. *Psychosom. Med.* **23**, 67-76.
- HALLER, S. VAN (1768). *Historia stirpium indigenarum Helvetiae*. Vol. II, p. 130.
- JONES, H. GWYNNE (1958). The status of inhibition in Hull's system: a theoretical revision. *Psychol. Rev.* **65**, 179-82.
- LACEY, J. I. (1950). Individual differences in somatic response patterns. *J. Comp. Physiol. Psychol.* **43**, 338-50.
- LYKKEN, D. T. (1957). A study of anxiety in the sociopathic personality. *J. Abnorm. Soc. Psychol.* **55**, 6-10.
- TONG, J. E. & MURPHY, I. C. (1960). A review of stress reactivity research in relation to psychopathology and psychopathic behaviour disorders. *J. Ment. Sci.* **106**, 1273-95.
- VOGEL, M. D. (1960). The relation of personality factors to GSR conditioning of alcoholics: an exploratory study. *Canad. J. Psychol.* **14**, 275-80.
- VOGEL, M. D. (1961). GSR conditioning and personality factors in alcoholics and normals. *J. Abnorm. Soc. Psychol.* **63**, 417-21.
- VORONIN, L. G., SOKOLOV, E. N. & BAO-KHUA (1959). Type features of the orienting response in man. *Voprosy Psikhologii*, **5**, 73-88.

(Manuscript received 14 July 1961)





## A THEORY OF THE SERIAL POSITION EFFECT

By EDWARD A. FEIGENBAUM

*University of California, Berkeley*

AND HERBERT A. SIMON

*Carnegie Institute of Technology*

The paper proposes a theory of the well-known serial position effect that makes quantitative predictions, acceptable by non-parametric tests, of the observed amount of bowing of the serial position curve. The theory, which stems from viewing the central nervous system as an information-processing system, is compared with the Lepley-Hull hypothesis and Atkinson's theory of the serial phenomena, and is shown to be more satisfactory than the older explanations.

### INTRODUCTION

Intraserial phenomena have been a major focus of interest in the study of serial learning. McCleoch (1942), for example, devoted 50 pages of his *Psychology of Human Learning* to such phenomena. And among intraserial phenomena, one of the most prominent is the serial position curve, depicting the relative number of errors made with the various syllables in a list while learning the list to some criterion.

McCrary & Hunter (1953) observed that if percentage of total errors is taken as the unit of measurement, then all the empirical serial position curves for lists of a given length are substantially identical. Earlier investigators, measuring number of errors, had concluded that relatively more errors occurred for the middle syllables when the lists were hard than when they were easy, more with slow learners than with fast learners, more with rapid presentation of syllables than with slow presentation, and so on. One can find in the literature numerous theoretical explanations of these differences.

The findings of McCrary & Hunter leave us in the embarrassing position of having explained phenomena that do not exist, i.e., the supposed differences in amount of the position effect, and of having failed to explain a striking uniformity that does exist—the substantial identity of curves derived under a variety of experimental conditions. McCrary & Hunter themselves reach the peculiar conclusion that a single principle can hardly be expected to account for uniformity of effect under diversity of conditions, and hence that some multiple-factor is needed to explain the outcome.

The thesis of this paper is the opposite one—that if a uniformity underlies experiments performed under a wide variety of conditions, this uniformity should be traceable to a single simple mechanism that is invariant under change of conditions. We shall propose such a model of the information processing activity of a subject as he organizes his learning effort in a serial learning task. The serial position effect will be shown to be a consequence of the information processing strategy postulated in the model; the model predicts both qualitatively and quantitatively the shape of the curve and the percentages reported by McCrary & Hunter and by others.

## SOME RELEVANT DATA

Before we state the theory of the serial position effect we shall review some important empirical findings on serial learning of nonsense syllables:

(1) Under the usual experimental conditions and with experienced subjects there is generally a characteristic curvilinear relation between number of errors to criterion for a given syllable and the serial position of that syllable in the list. The syllable with the largest number of errors is generally beyond the middle of the list, though this effect becomes less noticeable as the length of the list increases (Ribback & Underwood, 1950); the first syllable almost always exhibits the fewest errors (Hovland, 1938).

(2) McCrary & Hunter (1953) have shown that, for lists of a given number of syllables, all serial position curves obtained with the usual experimental procedures are virtually identical when errors are plotted on a percentage, rather than an absolute, basis. About the same degree of bowing is exhibited with nonsense syllables as with names, with massed as with distributed practice, with slow as with fast learners, with rapid as with slow presentation. Typical data for lists of 12 and 14 syllables are given in Tables 1 and 2.

(3) In spite of this uniformity under normal conditions, it is easy to produce large deviations from the characteristic curve. Such deviations can be produced in at least the following four ways: (a) by varying the difficulty of particular items in the list; (b) by introducing an item sharply distinguishable from those that precede it or follow it (McGeoch, 1942, p. 107); (c) by introducing distinguishable sublists within the main list (Wishner, Shipley & Jurvich, 1957); (d) by explicit instructions to the subjects (Krueger, 1932; Welch & Burnett, 1924). Not surprisingly, difficult items are learned with more errors than easy items in the same serial position; distinguishable items are learned with fewer errors; items that the subject is instructed to learn first are learned with fewer errors.

(4) For lists of a given length, average learning time per syllable is almost independent of: (a) the rate of presentation (though this result is not given explicitly by Hovland (1938), we have used his reported data to compute the average learning time per syllable; this constancy has been independently reported by Wilcoxon, Wilson & Wise (1961)); and (b) the order in which subjects are instructed to learn the items (though this result is not given explicitly by Krueger (1932), we have computed it from his data). Hence, number of trials to criterion is inversely proportional to seconds per syllable.

(5) Distribution of practice reduces the number of trials to criterion, but not sufficiently to compensate for the additional total time. The advantages of distribution, measured in trials to criterion, almost disappears when the presentation rate is as slow as four seconds per syllable (Hovland, 1938).

In the next section we shall propose a theory of the serial learning process that accounts quantitatively for the data mentioned in items 1, 2, and 4, above, and qualitatively for the observation of item 3. In the present paper we shall not discuss the effects of distribution of practice, since these effects almost certainly derive from mechanisms that go beyond the simple theory proposed here. We wish merely to observe that when time to criterion is taken as the measure of learning rate these effects are of rather small magnitude compared with those we shall consider.

*An information processing theory of serial learning*

We hypothesize that serial learning is an active, complex process involving the manipulation and storage of symbols by means of an interacting set of elementary information processes; and that these processes are qualitatively similar to those used in problem solving, concept formation, and other higher mental processes (Newell, Shaw & Simon, 1958*a*). Thus, we shall argue that the stimulus-response sequences postulated by S-R theory are simple only in surface appearance—that beneath them lies an iceberg of complex information processing activity.

We shall not defend this viewpoint in detail here although it has proved exceedingly fruitful in research in which we and our associates have been engaged (see, for example, Feigenbaum (1959); Newell & Simon (1961); Newell *et al.* (1958*b*)). We should like to offer three brief observations to persuade the reader that our conjecture does not entirely fly in the face of common sense or previous psychological observation. First, expectancy and mediation theories, like those of Tolman (Hilgard, 1956, pp. 185–221), or of Osgood (Hilgard, 1956, pp. 464–5), attribute as much complexity to the stimulus-response connexion as does our conjecture; what they fail to indicate is the nature of the mechanisms that might provide the complexity. Secondly, equally elaborate and more explicit mechanisms are postulated in concept-formation theories like those of E. J. Gibson (1940), and the recent one of Bruner, Goodnow & Austin (1956). Indeed, we shall see that one of our postulates involves a conception closely related to Bruner's notion of 'cognitive strain'. Thirdly, the time an experienced subject needs, per syllable, to memorize a list of a dozen nonsense syllables is of the order of thirty seconds. In comparison with the times required by familiar electronic systems for simple processes, this is an enormous time interval. It is large—by a factor of 500 or more—even in comparison with the 50 msec. or thereabouts required for the central processes in the simplest responses to stimuli. If a theory is to fill up this 30 sec. time interval in at all a plausible manner, it will have to attribute considerable complexity to the processes that take place. Feigenbaum (1959) reports on such a theory of verbal learning, dealing in a complete manner with discrimination learning, association learning, responding, etc. The theory predicts a variety of the phenomena of rote learning of nonsense syllables in serial and paired-associate learning tasks.

*Underlying assumptions of the model*

For the purposes of this paper, we shall not need to examine the elementary information processes in detail, for the shape of the serial position curve will prove to be independent of their micro-structure. This point is examined in detail by Feigenbaum (1959), where a distinction is drawn between macroprocesses of verbal learning and microprocesses. We require, instead, the assumption that in order for a connexion between a stimulus and a response to be formed, a certain (unspecified) sequence of elementary processes needs to be carried out, and that the execution of this sequence requires a definite interval of time, the length of the interval depending on the 'difficulty' of the task and other parameters of the experimental situation.

We suppose the information processing mechanism to be operating predominantly in a serial rather than a parallel manner—it is capable of doing only one, or a few



things at a time. The narrowness of the span of attention is a familiar aspect of conscious activity; we assume that it is also an attribute of the subconscious.

### *Information processing postulates*

The structure of the theory is embodied in four postulates about the processing mechanism.

*Postulate 1. Serial mechanism.* The central processing mechanism operates serially and is capable of doing only one thing at a time. Thus, if many things demand processing activity from the central processing mechanism, they must share the total processing time available. This means that the total time required to memorize a collection of items, when there is no interaction among them, will be the sum of the times of the individual items. (In serial learning of syllables there is, in fact, interaction among individual items; and total learning time increases more than proportionately with number of items. We will not be concerned with this point in the present paper because we are dealing not with total learning time or total errors, but with the relative number of errors made on different syllables in a list.)

*Postulate 2. Average unit processing time per syllable.* The fixation of an item on a serial list requires the execution of a sequence of information processes that requires, for a given set of experimental conditions, a definite amount of processing time per syllable. The time per syllable varies with the difficulty of the syllables, the length of list, the ability of the subject and other factors. In a well-known series of experiments by Hovland (1938), for example, it averaged approximately 30 sec.

*Postulate 3. Immediate memory.* There exists in the central processing mechanism an immediate memory of limited size capable of storing information temporarily; and all access to an item by the learning processes must be through the immediate memory. There is a great deal of experimental evidence to support the concept of an immediate memory. The evidence points to a span of immediate memory of about five or six symbols (Miller, 1956). We postulate that each symbol stored separately in the immediate memory must be a familiar, well-learned symbol. For unfamiliar nonsense syllable materials, the familiar symbols are the letters. Thus, for the three-letter nonsense syllables ordinarily used, we postulate that the immediate memory has the capacity to hold two syllables (six letters). This means that it will ordinarily hold at any moment one S-R pair being learned.

*Postulate 4. Anchor points.* In the absence of countervailing conditions—the nature of which will be specified presently—the information processing will be carried out in a relatively systematic and orderly way which will limit the demands that are placed on the small immediate memory. This postulate is related to the generalization, which Bruner and his associates (1956) have tested in certain concept-forming experiments, that subjects develop strategies for limiting the ‘cognitive strain’ involved in concept formation, and that these strategies involve handling newly acquired information in a systematic and orderly way.

We assume that subjects learning the syllables of a serial list will reduce the demands on memory by treating the ends of the list as ‘anchor points’, and by learning the syllables in an orderly sequence, starting from these anchor points and working toward the middle. This procedure reduces demands on memory because, at each stage of the learning task, the next syllable to be learned is readily identified

as being adjacent to a syllable that has already been learned. Thus, no special information about position in list needs to be remembered.

The idea of learning from anchor points is not new, though it does not seem to have been previously formalized. Woodworth (1938), for example, makes use of it in describing the process by which a list of digits is learned. Wishner *et al.* (1957) mention it in their discussion of the serial position curves obtained in their list-sublist experiment.

The first three postulates differ from the fourth in that the former describe built-in characteristics of the processing mechanism that are probably not learned or readily modified; while the latter describes a method of proceeding that is apparently habitual with most subjects, at least in our culture, but which is modifiable by experimental instructions, and by certain attention-directing stimuli.

It has been observed frequently that in serial memorization subjects not only develop associations between syllables, but also use various position cues and other cues. They learn, for example, that a particular syllable occurs in the early or in the late part of the list. (From this reliance on 'irrelevant' cues, one can develop an explanation for such phenomena as anticipatory errors and 'remote associations' that is much simpler than the usual one; but these topics would take us beyond the scope of our present task.) The use of position cues gives a unique status to the beginning and end of a serial list, for these items have the special property that they have no neighbours 'before' and 'after' respectively, i.e. the first item is always preceded, and the last item succeeded, by the intertrial activity. Once the items at the anchor points are memorized, the items contiguous to them become the first unlearned items 'after' or 'before' syllables already memorized; and so on, as the learning proceeds. More than this, the first two items are unique in that they represent the first S-R pair presented to the subject in the experiment. Thus, we can make out at least a plausible case that a learner can reduce the demands on immediate memory by memorizing in a more or less systematic fashion from the ends of the list toward the middle.

This postulate is sufficient to explain the bowed form of the serial position curve—although it says only a little more than the observed fact of the bowing. Its advantage over explanations like the Lepley-Hull hypothesis (which will be discussed later) is that it is not inconsistent with the ease (item 3, above) with which changes can be induced by the experimenter in the serial position curve.

In order to make quantitative predictions as to the amount of bowing that will be observed, we use the notion of anchor points to strengthen postulate 4, as follows:

*Postulate 4a. Processing sequence.* We postulate the following information processing strategy for organizing the serial learning task using anchor points: (a) the first two items presented in the experiment are learned first; (b) attention is next focused for learning on an item immediately adjacent to an anchor point. (In the ordinary serial list, this will be the third item or the last.) The probability that any specific item adjacent to an anchor point will be selected for learning next is  $1/p$ , where  $p$  is the number of anchor points (in the ordinary serial list,  $p = 2$ ). Thus, for example, the probability that the last item will be learned after the second item is 0.5; (c) attention is focused, and learning proceeds, item by item in this orderly fashion until the criterion trial is completed.

One can picture the subject building up over time an internal representation of the serial list he is learning. It will be seen, then, that the postulate specifies only a minimal amount of organizational activity: namely, the ability to add an item immediately after or immediately before an item already learned (or a 'special' stimulus like the intertrial interval). Our explorations with other processing strategies have shown that this strategy reduces greatly the information processing demands on the learner.

### *Predictions from the postulates*

The postulates describe a learning mechanism that memorizes serial lists in a prescribed way. This mechanism generates a serial error curve as it learns (i.e. some particular serial error curve is deduced as a consequence of the postulates). We wish to compute this serial error curve and compare it with the McCrary & Hunter curve.

Computer simulation is the most general and powerful method for doing this. We and others have used this method extensively in building theories of problem solving (Newell *et al.* 1958*a*), binary choice behaviour (Feldman, 1959), concept attainment (Hovland & Hunt, 1960), and other cognitive phenomena. It is described in detail elsewhere (Newell & Simon, 1961). Briefly the idea is this. The digital computer is a universal information processing device, capable of carrying out any precisely specified information process. Thus a computer can carry out exactly the information processing required by the postulates of the model. We programme the model on a computer, use it *qua* subject in verbal learning experiments (simulated inside the computer), observe the learning behaviour of the model, and thereby generate the consequences of the postulates in particular verbal learning situations. We have used this method in constructing and exploring an information processing theory of verbal learning (Feigenbaum, 1959). In particular, for the purposes of this paper, we have generated the serial error curve for a few simple serial learning experiments. We have done this in two different ways: first, following postulate 2, we introduced a unit processing time per syllable without specifying the microprocesses of the learning that take place during this time interval; secondly, we removed the latter artificiality and substituted the full complement of microprocesses postulated by the more complete theory.

For the particular case of the serial position curve, the postulates are simple enough so that there is no real need to employ computer simulation to generate the predictions. The postulates can be formalized in a simple mathematical model, from which the quantitative predictions can be generated. As this method is likely to be more familiar to the reader, we give the mathematical model in the Appendix, and present the serial error curves which it predicts (for lists of twelve and fourteen syllables) in Tables 1 and 2. The results obtained by the computer simulation technique are substantially identical (though slightly more discontinuous).

What can we specifically say about the fit? First, the ordinates of the first and last syllable of the predicted curves are in almost exact agreement with those of the empirical curves. Secondly, the syllable position of the peak of the predicted curves is substantially the same as that of the observed curves. Thirdly, the ordinates of the predicted curves and the empirical curves at each syllable position are very close, especially in the critical first and last third of each list, where very good agreement is



important to any claims about goodness-of-fit. Furthermore, this fit was obtained without any arbitrary parameters, other than the specification of the sequence in which incoming syllables are processed.

The goodness-of-fit of the observed frequency distribution to the predicted distribution was tested by the Kolmogorov-Smirnov test of association. The test accepted the null hypothesis at the 99% level of significance.

Table 1. Table showing percentage of total errors made during acquisition at each syllable position of a 14-item serial list of nonsense syllables predicted and observed

	Syllable position													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Predicted	0.95	1.9	4.7	5.6	8.1	8.9	10.5	10.8	10.8	10.5	8.9	8.1	5.6	4.7
Observed*	1.0	3.5	4.3	6.0	8.0	8.9	9.2	10.0	9.5	10.6	8.8	8.9	7.2	4.0

\* These values are approximate.

The data were taken from fig. 4 of McCrary & Hunter (1953, p. 133).

Table 2. Table showing percentage of total errors made during acquisition at each syllable position of a 12-item serial list of nonsense syllables predicted and observed

	Syllable position											
	1	2	3	4	5	6	7	8	9	10	11	12
Predicted	1.3	2.6	6.3	7.5	10.5	11.3	12.5	12.5	11.3	10.5	7.5	6.3
Observed*	1.5	2.3	4.5	7.0	10.3	11.6	14.0	12.4	11.0	10.0	8.5	7.0

\* These values are approximate values for the median percentages at each position for the four curves presented in fig. 2 of McCrary & Hunter (1953, p. 132).

### ELABORATION AND DISCUSSION

In this section, we wish to compare the predictions of our information processing theory with those derived from the Lepley-Hull hypothesis, and extend our predictions to two important experiments, one of which was published after we had specified our model.

#### (1) The Lepley-Hull hypothesis

There have been few attempts to account for the serial position effect in quantitative terms. Hull *et al.* (1940) attempted to do so on the assumption of some inhibitory processes, or intralist 'interference'. Atkinson (1957), drawing on statistical learning theory, has exhibited a stochastic process which generates a curve of the general shape of the serial position curve. We shall discuss Hull's results in some detail, and comment briefly on Atkinson's.

Although Hull's equations provide a good fit to the empirical data, this fit is not a convincing test of the Lepley-Hull theory for the following reasons: the reaction threshold, the ratio of inhibitory potential to excitatory potential per trial, and the remoteness reduction factor. These are used to fit the serial error curve, or, more precisely, the curve of number of repetitions to reaction threshold (see Hull *et al.* (1940, pp. 103-7)). Hull fits the theoretical curve by passing it through three points of the

empirical curve. Since the empirical data form a relatively smooth, bow-shaped curve, it is not surprising that a three-parameter curve can be made to fit them closely; an equally good approximation can be obtained by fitting a parabola empirically to the data.

This means that Hull's hypothesis will fit almost any data (provided the serial position curve has the characteristic bowed pattern), and hence is almost impossible to disprove from the data. It is therefore an exceedingly weak hypothesis. By the same token—because of the three free parameters—Hull's theory does not predict the constancy on a percentage scale observed by McCrary & Hunter.

Conversely, given the constancy observed by McCrary & Hunter, we can draw certain conclusions from Hull's theory regarding the growth of excitatory and inhibitory potential and the reaction threshold. For example, it can be shown by an examination of Hull's equations that the ratio of the increment of inhibitory potential per trial to the increment of excitatory potential must be a constant (independent, for example, of intra-list similarity). See Hull *et al.* (1940, pp. 104–5).

The McCrary & Hunter result implies that the  $R$ s are related: they are proportional to each other. The variables  $q_i$  are homogeneous of degree two in  $R$ , and the  $D_i$  are homogeneous of degree one in  $R$ . The  $J$ s are homogeneous of degree zero in  $R$ . By equation (3) of p. 104,  $\Delta e/\Delta k$  is homogeneous of degree zero in  $R$ , hence a constant.

This is surprising and contrary to the whole spirit of the Lepley-Hull hypothesis. For we would expect that with high intra-list similarity the inhibitory potential would rise more rapidly than with low intra-list similarity. On the contrary, if Hull's model is correct the only parameter that changes as lists become more difficult to learn is the ratio of the threshold to the increment of excitatory potential per trial (by equation 2, p. 104, of Hull *et al.* (1940)). Finally, the Lepley-Hull hypothesis does not explain how a subject can voluntarily or through a shift in his attention greatly alter the shape of the curve.

There are four reasons, therefore, why Hull's mathematical model for the serial position effect is unsatisfactory: since it contains three adjustable parameters its predictions are very weak; it does not predict the constancy in the percentage-error curve; this constancy hardly seems compatible with the mechanism assumed as a basis for the model; and finally, the model is difficult to reconcile with well-known attention-shift and set-change phenomena.

The preceding discussion of the curve-fitting aspects of Hull's equations applies also to Atkinson's equations. Atkinson (1957) has available four free parameters. He estimates these parameters from data for an 18-syllable list, and uses these estimated values to make predictions for lists of 8 and 13 syllables. However, a careful examination of Atkinson's equations shows that even after the parameters have been estimated, there are enough degrees of freedom left in the system almost to insure a reasonably good fit to the other curves. Thus, his theory suffers the same infirmity that we have pointed out in Hull's.

Furthermore, to make workable the difficult mathematics of the stochastic process, Atkinson has had to introduce a number of very constraining assumptions. His equations hold only for serial lists of highly dissimilar words which are familiar and easily pronounced; the presentation must be at a moderate rate, with a long interval

provided at the conclusion of each trial. Yet the same bowed curve is obtained empirically when these conditions are not met as when they are. Finally, as in Hull's theory, there is difficulty in predicting the constancy observed by McCrary & Hunter. (Given a set of parameter values, Atkinson's theory does not predict the constancy over the experimental conditions reported by McCrary & Hunter. On the other hand, if one admits that these values may change from situation to situation, then one must re-estimate them for each experimental condition, and therein the theory loses much of its power.

## (2) *The two-part list*

A simple extension of ordinary nonsense syllable experiments is to differentiate the first half of the list from the second half by printing the former in one colour (say black) and the latter in another colour (say red), dividing the total list perceptually into two smaller sublists.

What will be the shape of the serial position curve? One can predict this from Hull's theory by the unsatisfying procedure of assuming that the total list is learned as two separate sublists, and by fitting each sublist with a three-parameter curve such as we have discussed previously. The total fit will have six free parameters.

We should like to be able to predict the shape of the curve from the theory we have already presented. There are two important issues involved in making such a prediction:

(i) Consider those subjects who perceive the total list as being constructed of two sublists. One possible reasonable strategy for dealing with the learning task is to use the end points of each sublist as anchor points in the type of learning process described earlier. Another plausible strategy is to use as anchor points the end of the total list and one point in the centre to identify the point of bifurcation, say the first red syllable.

In making a prediction of the serial position curve for the two-part list, we have assumed simply that of those subjects who perceived the list as being two sublists, one half used the first strategy and the other half used the second strategy. The assumption is, of course, a relatively crude one, but the prediction is not very sensitive to the actual percentages assumed. Alternatively, we could have estimated the percentage from the data.

(ii) In the experiment which we shall discuss shortly, we have no way of knowing precisely how many subjects perceived the task as one of memorizing two sublists and how many perceived it as learning one long serial list. In the absence of this knowledge, we can estimate these percentages from the observed ordinate of the first red syllable and weight our prediction essentially 'fits' our predicted curve to the using this estimate. This procedure essentially 'fits' our predicted curve to the empirical curve at one point, the 'break' between black and red syllables. But it guarantees nothing about the quality of the prediction at the other points.

As part of a larger experiment, Wishner *et al.* (1957) performed an experiment with a two-part list. They had an experimental group memorize lists of 14 syllables, half of which were printed in black capitals, the other half in red lower case letters. The experimental group was told that the object of the experiment was to discover how people learned two lists simultaneously.



In Table 3 the predicted values for the percentage of errors at each syllable position are compared with the observed values.

The predictions were generated by the methods previously described. Three anchor points were used. In the mathematical treatment the three-anchor-point predictions were corrected by a factor which insured an exact fit of the ordinate of syllable 9 (the middle anchor point). In the computer simulation method, whole-list and sublist strategies were both run and the predicted ordinates averaged in a weighted fashion such that the ordinate of syllable 9 fitted exactly. What this procedure comes down to is the assumption that approximately two-thirds of the subjects learned the list as two sublists and approximately one-third learned it as a single long list (note that these fractions were not assumed *a priori* but were obtained by working backwards from the observed ordinate of the ninth syllable).

Table 3. *Table showing percentage of total errors, predicted and observed, made at each syllable position during acquisition, for a two-colour, 14-item serial list of nonsense syllables*

	Syllable position													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Predicted	1.9	3.1	4.7	6.7	9.1	10.1	8.9	5.9	6.0	8.3	8.7	9.8	9.9	5.5
Observed*	2.2	4.3	6.5	7.3	8.2	8.5	8.1	6.0	6.0	9.0	9.5	9.9	8.5	6.0

\* These values are approximate, and were derived from data taken from Wishnor *et al.* (1957, p. 260).

The agreement at all syllable positions is very close. As contrasted with our other predictions, in this one we had available the one free parameter already mentioned. The goodness-of-fit of the observed frequency distribution to the predicted distribution was tested by the Kolmogorov-Smirnov test of association. The test accepted the null hypothesis at the 99 % level of significance.

### (3) *The experiments of Krueger*

We turn now to some important experiments, the results of which were alluded to previously, and which have important implications for the information processing theory.

In a well-known series of experiments, Krueger (1932) presented various kinds of lists of 'easy' and 'hard' paired nouns to subjects who either received instructions to learn the list in some specified order or no instructions at all. These studies demonstrated that the order in which the various items were learned was influenced markedly by the instructions given to the subjects. As McGeoch puts it (1942, p. 102). 'The relation between rate of learning and position in the series is, then, a function of the direction of the subject's effort or attention'. As we have indicated, this is entirely consistent with the information processing theory, which regards particular learning sequences as 'strategies' for dealing with the learning problem—as adaptive response to task.

What about our more specific hypothesis that in the usual serial learning experiment the 'end points' of the list will be taken as anchor points in the learning process? Krueger's experiments showed that subjects given *no* instructions produced

essentially the same serial position curves as those subjects who were instructed to learn the *ends* of the list first.

Because the fixation of an item requires a fixed amount of processing time, and because the sequence of learning is considered a 'strategy' and not a built-in characteristic of the learning process, our theory predicts that the total number of syllables learned will be proportional to the total learning time and independent of the order of learning. On this point Krueger reports, 'When the attention given is constant, the total amounts learned are the same, irrespective of whether this effort is directed to the beginning, center, or the final sections of the unit which is to be memorized' (1932, p. 527).

Although we assume that there is a constant fixation time associated with each particular item on a list, items of different kinds (e.g. 'easy' items against 'difficult' items) will have different processing times. The theory we have proposed predicts that the total learning time will be the sum of these processing times per syllable, and as such will be independent of the order in which the syllables are learned. Confirming this prediction, Krueger reports, 'When materials of unequal difficulty appear within the same unit to be mastered, the total number of trials required to memorize the unit is approximately the same *whether attention is given at first to the more difficult or the easier sections of the unit*' (1932, p. 527).

This is consistent with the McCrary & Hunter results. If one plots the McCrary & Hunter curve by ordering the abscissa values not by serial position in the list but by the apparent order in which the syllables were learned, the ordinates lie on a straight line. This, of course, is in exact agreement with our model. Recent additional information on this phenomenon was obtained by Jensen (1962) for the learning of nonsense figures.

Thus, Krueger's experiments, though they were performed with serial lists of paired nouns rather than nonsense syllables, demonstrate (a) that the serial position curve can be 'shaped' by the experimenter with suitable instructions to subjects, so that the order of learning syllables is itself a learned response; (b) that the total amount of material learned within a given time is independent of the order in which various items, sometimes heterogeneous with respect to difficulty, are learned.

#### SOME RECENT RESULTS

Subsequent to the specification of the model proposed in this paper, some important new experiments have been published on the effect of replacing syllables on a list during learning. Rock (1957) used the following procedure on his experimental groups: on each test trial, those syllables incorrectly responded to by the subject were removed and new syllables were substituted in their place for the next learning trial. Rock found no impairment of the rate of learning for the experimental groups (as compared with the control groups). This important result casts further doubt on Hull's incremental build-up hypothesis. Criticism of Rock's technique led Estes, Hopkins & Crothers (1960) to replicate and extend Rock's experiment, but their results substantiate Rock's.

Estes *et al.* say of these experiments, 'No hitherto published theory with which we are familiar gives a reasonable amount of our principal findings' (1960, p. 338). The

information processing theory we have proposed here predicts the Rock result. A computer simulation of the Rock experiment using the information processing model generated behaviour substantially identical with that reported by Rock. In terms of our theory (postulate 4a) the explanation, of course, is that items on a list are learned one at a time in the processing sequence. Items presented when attention is focused on some other particular item are simply ignored by the learner, and are picked up on a later trial, as determined by the processing sequence. Hence, no time is lost by the learner if the experimenter replaces an item that has not yet been processed.

#### CONCLUDING REMARKS

In this paper we have surveyed the principal known facts about the shape of the serial position curve in serial learning of nonsense syllables by the anticipation method. We have examined the Lepley-Hull hypothesis as an explanation for the shape of the curve, and have concluded that the hypothesis is unsatisfactory. We have proposed an alternative hypothesis formulated in terms of information processes. We have shown that the hypothesis not only predicts the constancy of the serial position curve when the ordinates are plotted in percentage terms, but also predicts the quantitative values of the ordinates. Since the hypothesis allows no free parameters, its success in fitting the observed data provides rather persuasive evidence for its validity.

The information processing hypothesis is built on the following assumptions:

- (1) that the brain is a serial processing mechanism with a limited span of processing attention;
- (2) that the fixation of an item uses up a definite amount of processing time;
- (3) that there is a small immediate memory which holds information to be processed;
- (4) that the subject employs a relatively orderly and systematic method for organizing the learning task, using items with features of uniqueness as anchor points.

In this paper, we have offered no explanation of the fixation process itself, i.e. we have talked not at all about what occurs during the processing time assumed in item (2) above.

We are indebted to our colleague Allen Newell for numerous helpful discussions about this project, and to the Ford Foundation for financial assistance that made it possible.

#### APPENDIX

Given the unit processing time per syllable, one can, from the postulates, compute the average time after the beginning of a learning experiment that will pass before any specified syllable is learned. This time, in turn, determines uniquely the number of errors that will be made with that syllable. While the actual number of errors will be a function of the unit processing time, the percentage that this number represents of the expected total errors is independent of the unit processing time.

The numerical estimates given in the text tables were obtained as follows: By the postulates, the syllables will be learned in an orderly sequence, each syllable requiring a certain processing time, say  $k$ . Each syllable can be identified by its serial order,  $i$ , in the list as presented by the experimenter, and also by the order,  $r$ , in which it is learned by the subject. Since learning takes



place from both ends of the list, these two orders will not, in general, be identical. Thus,  $s_i$ , the  $i$ th syllable in order of presentation, may be the same syllable as  $s'_r$ , the  $r$ th syllable in order of learning. (Technically, the list of syllables in order of learning is a permutation of the original list.) Let  $T'_r$  be the time that elapses before the first successful response to syllable  $s'_r$ —that is, until the  $r$ th syllable is learned. Then  $T'_r = kr$ . The number of errors,  $W'_r$ , the subject will make on the  $r$ th syllable is equal to the number of learning trials prior to the trial on which that syllable is learned and will be proportional to  $r$ :

$$W'_r = mr, \quad (1)$$

where  $m$  is a proportionality constant, equal to  $k$  divided by the time per trial.

The numerical value of  $m$  is a function, of course, of the difficulty of the items, and the rate of presentation. However, we are only concerned with the fraction of total errors made on a given syllable, and this fraction is clearly independent of  $m$ . For let  $W$  be the total number of errors, summed over all syllables, and  $w'_r = W'_r/W$ , the fraction of total errors made on the  $r$ th syllable.

$$w'_r = mr / \sum_{r=1}^n mr = r / \sum_{r=1}^n r = \frac{2r}{n(n+1)}, \quad (2)$$

where  $n$  is the number of syllables in the list.

Suppose, for example, that we are dealing with a list of twelve syllables, then  $\sum_{r=1}^{12} r = 78$ . Hence, the fraction of total errors that will be made on the 4th syllable learned will be  $4/78 = 0.051$ .

Now, to obtain the serial position curve, we need merely to relabel the syllables from the order in which they are learned to their order of presentation. That is, if  $r_i$  is the rank, in order of learning, of the  $i$ th syllable in order of presentation, then the fraction of total errors for the  $i$ th syllable will be simply:

$$w_i = w'_{r_i} = r_i / \sum_r r_i. \quad (3)$$

To apply this result, we must calculate the rank,  $r_i$ , of the  $i$ th syllable in order of presentation, as determined by postulate 4a. We assume (postulate 3) that the immediate memory capacity is two syllables and for simplicity in the calculation shall assume that the items are picked up pairwise in the processing sequence, and stored in the immediate memory for learning. The first two syllables in the list will be learned first (will have rank,  $r = 1$ , and  $r = 2$ , respectively), followed either by the last two syllables on the list, or by syllables three and four, each with probability one-half. The result is that in a list of 12 syllables the third syllable, for example, will have a probability of one-half of being the third syllable in order of learning, a probability of one-quarter of being the fifth syllable, a probability of one-eighth of being the seventh, and of one-sixteenth of being ninth or eleventh. Averaging these ranks, weighted by their respective probabilities, we find that the average rank of the third is  $r_3 = 4.875$ . The fraction of total errors on the third syllable will then be  $w_3 = 4.875/78 = 0.063$  (see Table 2). All the other predicted values in the tables were computed in the same way.

The fact that the third syllable has zero probability of learned fourth, sixth, eighth, etc., is artificially introduced by the calculation simplification introduced above of handling the syllables in pairs. It does not materially affect the serial position curve prediction, as is shown by the fact that the computer simulation (which does not use the pairwise learning simplification) generated the same serial position curve prediction.

## REFERENCES

- ATKINSON, R. C. (1957). A stochastic model for rote serial learning. *Psychometrika*, **22**, 87-95.  
 BRUNER, J. S., GOODNOW, J. J. & AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.  
 ESTES, W. K., HOPKINS, B. L. & CROTHERS, E. J. (1960). All-or-none and conservation effects in the learning and retention of paired associates. *J. Exp. Psychol.* **6**, 329-39.  
 FEIGENBAUM, E. (1959). *An Information Processing Theory of Verbal Learning*. The RAND Corporation, Paper P-1817.

- FELDMAN, J. (1959). Analysis of Behaviour in Two Choice Situations. Unpublished doctoral dissertation, Carnegie Institute of Technology.
- GIBSON, E. J. (1940). A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.* **47**, 196-229.
- HILGARD, E. R. (1956). *Theories of Learning* (2nd ed.). New York: Appleton-Century-Crofts.
- HOVLAND, C. I. (1938). Experimental studies in rote learning theory. III. Distribution of practice with varying speeds of syllable presentation. *J. Exp. Psychol.* **23**, 172-90.
- HOVLAND, C. I. (1951). Human learning and retention. In S. S. Stevens (ed.), *Handbook of Experimental Psychology*. New York: Wiley.
- HOVLAND, C. I. & HUNT, E. B. (1960). The computer simulation of concept attainment. *Behavioral Science*, **5**, 265-7.
- HULL, C. L. *et al.* (1940). *Mathematico-deductive Theory of Rote Learning*. New Haven: Yale University Press.
- JENSEN, A. R. (1962). Is the serial position curve invariant? *Brit. J. Psychol.* **53**, 159-66.
- KRUEGER, W. C. F. (1932). Learning during directed attention. *J. Exp. Psychol.* **15**, 517-27.
- MCCRARY, J. W. & HUNTER, W. S. (1953). Serial position curves in verbal learning. *Science*, **117**, 131-4.
- MCGEOCH, J. A. (1942). *The Psychology of Human Learning*. New York: Longmans, Green.
- MILLER, G. A. (1956). The magical number seven, plus or minus two; some limits on our capacity for processing information. *Psychol. Rev.* **63**, 81-97.
- NEWELL, A., SHAW, J. C. & SIMON, H. A. (1958a). The elements of a theory of human problem solving. *Psychol. Rev.* **65**, 151-6.
- NEWELL, A., SHAW, J. C. & SIMON, H. A. (1958b). *The Processes of Creative Thinking*. The RAND Corporation, Paper P-1320.
- NEWELL, A. & SIMON, H. A. (1961). Computer simulation of human thinking. *Science*, **134**, 2011-17.
- RIBBACK, A. & UNDERWOOD, B. V. (1950). An empirical explanation of the skewness of the bowed serial position curve. *J. Exp. Psychol.* **40**, 329-35.
- ROCK, I. (1957). The role of repetition in associative learning. *Amer. J. Psychol.* **70**, 186-93.
- WELCH, G. B. & BURNETT, C. T. (1924). Is primacy a factor in association-formation? *Amer. J. Psychol.* **35**, 396-401.
- WILCOXON, H. C., WILSON, W. R. & WISE, D. A. (1961). Paired-associate learning as a function of percentage of occurrence of response members and other factors. *J. Exp. Psychol.* **61**, 283-9.
- WISHNER, J., SHIPLEY, T. E., Jr. & JURVICH, M. S. (1957). The serial position curve as a function of organization. *Amer. J. Psychol.* **70**, 258-62.
- WOODWORTH, R. S. (1938). *Experimental Psychology*. New York: Holt.

(Manuscript received 27 June 1959; revised 11 July 1961)

## PERCEPTUAL-MOTOR TRANSFER IN IMBECILES: A SECOND SERIES OF EXPERIMENTS

BY A. D. B. CLARKE AND MARGARET COOKSON

*Psychology Department, The Manor Hospital, Epsom, Surrey*

Earlier experiments tested a hypothesis derived from Hebb, that perceptual-motor transfer would be negatively correlated with age in imbecile subjects of 9, 17 and 23 years. On four tasks this hypothesis was strongly confirmed, with the youngest showing massive transfer of training compared with the older subjects, and almost reaching adult level. The present series of experiments included tasks of greater complexity. An unexpected finding was that the younger subjects were able, after 6 months of non-practice, to learn a more difficult sorting task more easily than the easier one 6 months earlier, thus showing the retention of learning set and improved perceptual and conceptual discrimination. Another experiment with the complex Minnesota formboards again showed greater transfer in children than in adults; here an initial average gap of 640 sec. between the performance of children and that of adults was narrowed to 45 sec. after 32 trials. Finally, after a year of non-practice, adolescent and adult subjects were able to carry out a sorting task very much better than previously, indicating the long-term effect of earlier learning. Implications for theory and practice are discussed.

### I. INTRODUCTION

An earlier paper (Clarke & Blakemore, 1961) investigated the hypothesis that transfer of training was inversely proportional to age, using imbecile subjects of different ages. This hypothesis was strongly confirmed. It was found that 9-year-old imbecile subjects showed massive transfer on a number of perceptual-motor tasks, as compared with those aged 17 and 23. This was sufficient to raise the efficiency level of the children almost to that of adults of the same type, and it was concluded that this transfer involved both learning set and a sharpened perceptual and conceptual discrimination.

The present series of experiments was designed to investigate further the learning processes of the same 9-year-old subjects, employing the same type of sorting tasks as those used in the previous research. Experiments I and II (card sorting and type-writer key sorting) used designs of much greater complexity than previously. A third, entirely new, experiment was also necessary, and a fourth checked an inference from the second, namely that, using older subjects, learning would be retained over a long period of non-practice. There were two reasons for these experiments: first, we were interested in finding out whether the same pattern of transfer occurred in a much more difficult task; and secondly, by using more difficult problems and thereby leaving more room for improvement, we wanted to know whether the young imbeciles would once again almost reach adult level.

### II. EXPERIMENT I. DESIGN SORTING

The initial experiment used a much more difficult version of the sorting of cards bearing abstract designs. In the original (1961) experiment the five designs in each of the parallel forms, though abstract, were easily distinguishable. One of the newer sets of designs is shown below in Fig. 1.



The subject was required to sort out fifty cards, measuring 4 in. square, into five labelled trays. On initial trials the subjects took from 10 to 15 min. per trial and although on the sixth trial the time had been reduced to an average of 6 min., this mirrored only the increased speed with which the subjects were now able to flip the cards from one tray to the next, awaiting the decision 'good' or 'no' from the experimenter, and the error score remained high.

Only one child from the group had begun to make one consistently correct placement (of design C) while the remainder of the group were becoming bored with the task. It seemed that the main problems experienced by these subjects arose, first, from the great similarity of the designs which differed only in the orientation of one

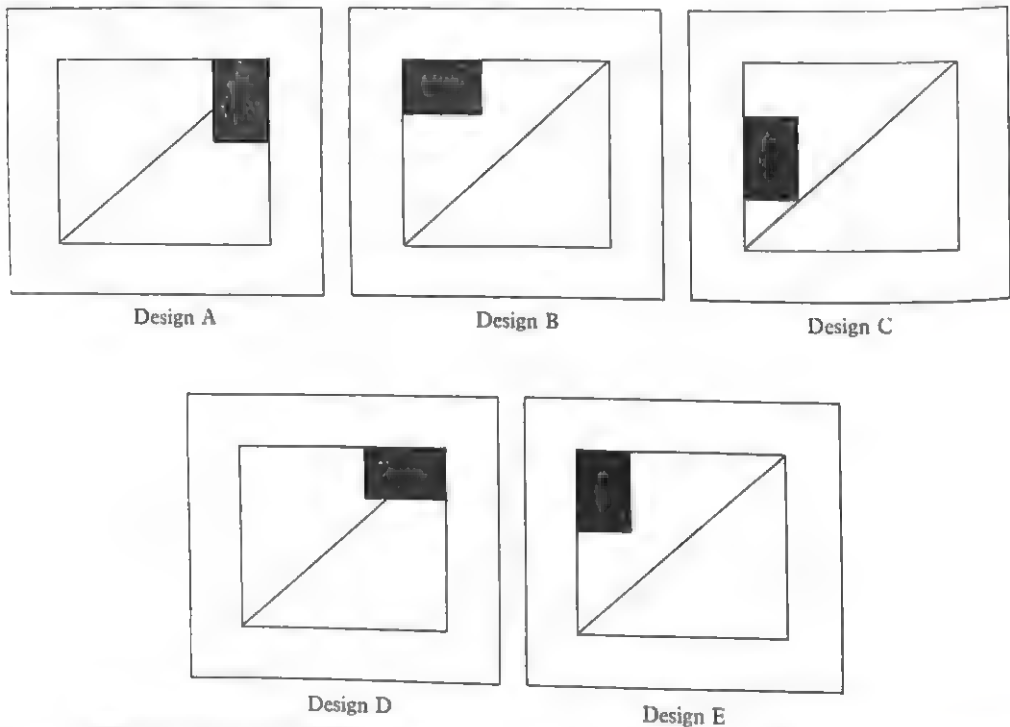


Fig. 1. Five designs to each of which ten cards had to be matched from a pack in which they had been placed in random order.

component, and, secondly, from the use of diagonals and asymmetry which maximised their perceptual difficulties; at a much higher level the same type of confusion and misorientation has been shown by Shapiro (1951, 1952, 1953) with brain-damaged patients. With time limits imposed, for administrative reasons, upon the duration of the experiment, it was therefore decided to abandon it and proceed with the typewriter key-top sorting task which had the advantage of being considerably shorter and in which, therefore, motivation was less of a problem.

Subsequently, a pilot experiment investigated the principles of teaching this type of task, using the material shown in Fig. 1, and two of the children who had previously been unable to learn it. The approach suggested by Itard (an early nineteenth-century forerunner of Hebb, who gave evidence for the dependence of learning upon earlier learning) seemed obviously appropriate. Itard had shown that if a learning barrier were reached, one needed to analyse the processes responsible,

revert to an earlier, simpler level, strengthen the weak processes by practice, and then return to the barrier and surmount it. The first principle was to reduce the number of cards from 50 to not more than 30, and to start by practising discrimination of only two designs (C and D) with 10 cards each. The next step, when this was mastered, was to increase the discriminations to three (A, C and D), still with ten cards each, and then to four with seven cards each (A, B, C, and D). Finally, the full number of discriminations, five (with six cards each), represented the complete task. One child improved steadily and after only a few trials daily over 8 days could achieve final performance with only three or four errors. The other child was unable to achieve good discrimination of two designs, and teaching was therefore abandoned for the time being. Practice over a longer period and with greater delays at each new step would be needed in his case if success were to be achieved. Nevertheless, this pilot experiment gave general support in the one case to the principles of teaching evolved by Itard, and suggested that these apparently insoluble problems could eventually be learned.

### III. EXPERIMENT II. KEY-TOP SORTING

The second task used was the sorting of twenty-five typewriter key-tops according to their symbols. This task, like the others, was split into two approximately equal forms (A and B), each consisting of five different symbols with five keys bearing each symbol. The subject was required to sort each set of twenty-five keys into five labelled tins.

Fig. 2 shows (a) the symbols of each form of the original sorting task (carried out in March, 1960) and (b) the more difficult sorting task (carried out in October, 1960).

It will be noted that the later problems were planned to be confusing to the subject, by including within each group of five keys two pairs of symbols which differed from each other only in minor respects.

Exactly the same procedure was used as in the earlier (1961) experiment, that is, each subject was given two trials per day for 1 week on each task. As before, the order in which the two parallel forms were administered was randomized in order to cancel any slight difference in their difficulty. The subjects were rewarded with sweets, irrespective of performance, and, as usual, motivation appeared high.

The most outstanding result of this part of the experiment was its failure to answer the first question above—whether in fact the same pattern of transfer would occur in a much more difficult task. We had expected, as implied earlier, a learning curve starting on this more difficult task a long way above the learning curve of the earlier experiment in March, 1960. Instead, as will be seen in Fig. 3, the new results were significantly better than those of the earlier experiment, and suggested that the more difficult task was in fact easier.

It remained just possible, however (although very unlikely), that the task chosen, which normal people considered more difficult than the earlier one, was in fact not so for the imbecile, to whom all symbols, old and new, were unfamiliar and intrinsically meaningless. Thus, to eliminate this possibility, we took a further 16 imbecile subjects and gave four groups of four the earlier and the present task in alternated order. These were studied merely on the initial trial of each series. The results showed that on the average the initial trial of the new task was the more difficult by one minute

(the mean of the initial times on the old task being 113 sec. and on the new, 173 sec.). Hence the new task was, in fact, much more difficult than the earlier one.

Our main finding is thus that the 9-year-old imbeciles were able in September 1960 to do a much more difficult task more easily than they did an easier task in March 1960. It seems clear that two processes were at work. First, some retention of learning set: for example, the children knew what was required of them initially in October without the need for instruction. The second factor seemed to be a sharpened perceptual, and indeed conceptual, discrimination during the experiment. This could not possibly be attributed to maturation over six months in persons of this type.

A (a)



B



A (b)



B



Fig. 2. The sorting of plastic typewriter keys. This figure (a) shows the two easier parallel forms learned in March 1960, and (b) the two more difficult parallel forms learned by the same subjects in October 1960.

#### IV. EXPERIMENT III. MINNESOTA FORMBOARDS

In a further quest for a very difficult task to replace the too difficult material employed in Experiment I, it was decided to use the four formboards of the Minnesota Spatial Relations Test. A good deal of work using adult imbeciles has, of course, been carried out with these (e.g. Tizard & Loos, 1954).

This test consists of four formboards, A, B, C and D, each containing 58 holes with 58 corresponding pieces set out in a constant position in front of each board. The subject had to place each shape into its correct hole as rapidly as possible. There were 58 shapes common to boards A and B and a further 58 common to boards C and D. As implied above, our use of the formboards was not as a *test* but as a *learning task*.



We used as control data material gathered on 5 adult imbeciles by Tizard & Loos earlier, and each 9-year-old subject was matched with an adult of approximately similar grade. Using exactly the same procedure as in the 1954 experiment, each subject carried out eight learning trials per board in the same order as had his control. This task, because of its size and complexity, and, therefore, initially its duration (which on first trials averaged 30 min., with a range of 18-50 min.) was the most

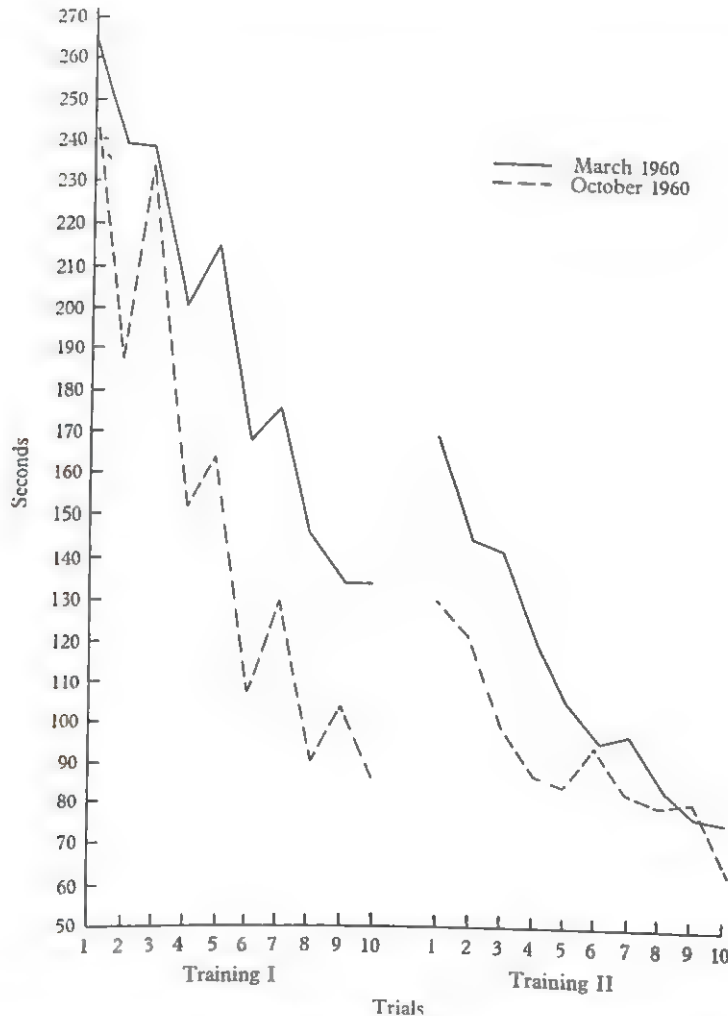


Fig. 3. The average performance of 5 9-year-old imbecile subjects in sorting typewriter keys in March 1960 (easier version, Fig. 2 (a)) and in October 1960 (more difficult version, Fig. 2 (b)).

ambitious yet attempted with imbecile children, and we encountered, for the first time, motivational difficulties, having in fact to drop one little girl, and replace her by another of the same grade. Boredom also seemed to be a factor in that, rather than being put off by a new board, most subjects showed pleasure and excitement at tackling something new. Two subjects actually reduced their times from the final trials on Board II to the first trial on Board III, involving a change both of boards and of shapes. Fig. 4 shows the main results.

It will be noted that the same pattern of findings emerged from this as from the four earlier experiments reported by Clarke & Blakemore (1961). The starting point of the children was extremely slow. Thereafter steady but moderately rapid improvement occurred and at the end there was an average of only 45 sec. between the children and the adults, whereas at the start the discrepancy had averaged 640 sec. Distractibility is of course a common feature in brain-damaged children and this was most prominent in the early trials of this experiment. However, as learning progressed, distractible behaviour decreased and final trials were remarkable for their demonstration of the imbeciles' uninterrupted concentration for an average of about five minutes per trial.

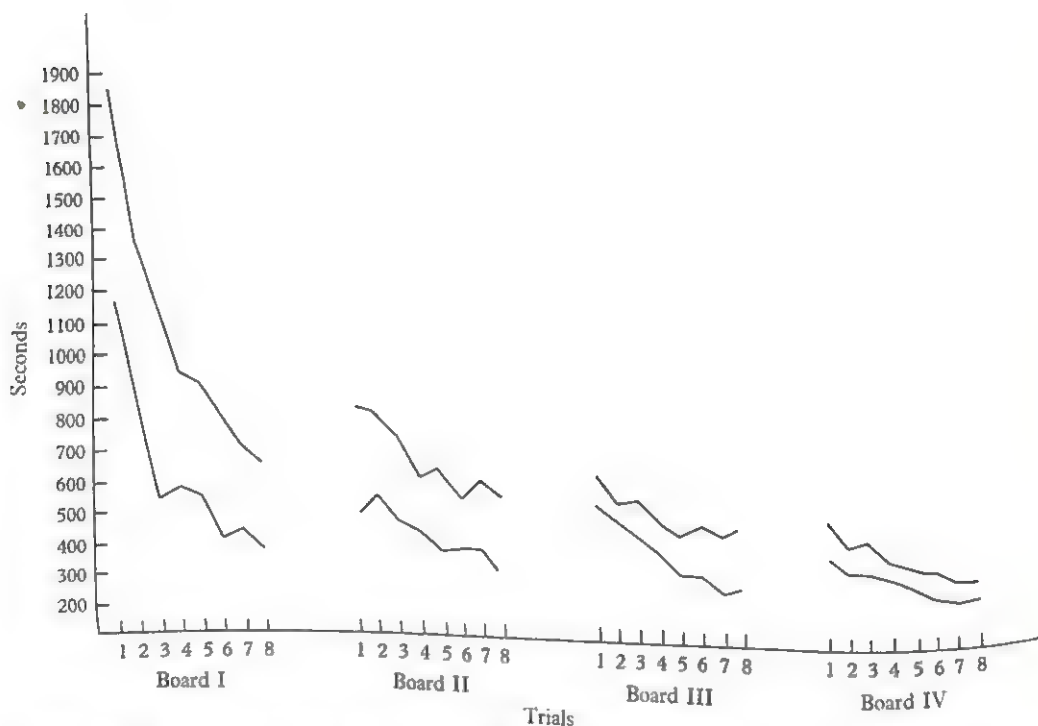


Fig. 4. The Minnesota Spatial Relations Test. This figure shows the average performance of 5 9-year-old imbecile children (upper curves), and 5 adult imbeciles (lower curves) in the learning of the four formboards.

It may also be worth noting that three of the children spontaneously named some of the shapes ('Christmas tree', 'car', 'football', 'bed', etc.) and this appeared to facilitate their correct placement. The use of verbal concepts in this way was unexpected in imbecile children in view of Luria's well-known work (1961).

#### V. EXPERIMENT IV. KEY-TOP SORTING

It will be recalled that in the second experiment, the 9-year-old imbeciles were able to do a more difficult typewriter key sorting task more easily than an easier one performed 6 months earlier, showing retention of learning set and an improved perceptual and conceptual discrimination.

The present experiment was set up to test the prediction that the adolescent and adult subjects would show the same type of improvement on the same sorting task (as opposed to a different one) even one year after their first experience of it. Accordingly, the 6 adolescent and young adult imbeciles available from the original ten (the remainder being in quarantine) were retested on the same material. Naturally, the two equivalent tasks were presented to each subject in the same order as a year previously; hence he did not begin the experiment on the last task he had learned but the first.

It was decided to weight the scales against the prediction by (a) omitting the pre-test given by Clarke & Blakemore (1961); and (b) by giving no instructions whatsoever, other than: 'You have done this before; do it as quickly as you can'. It was

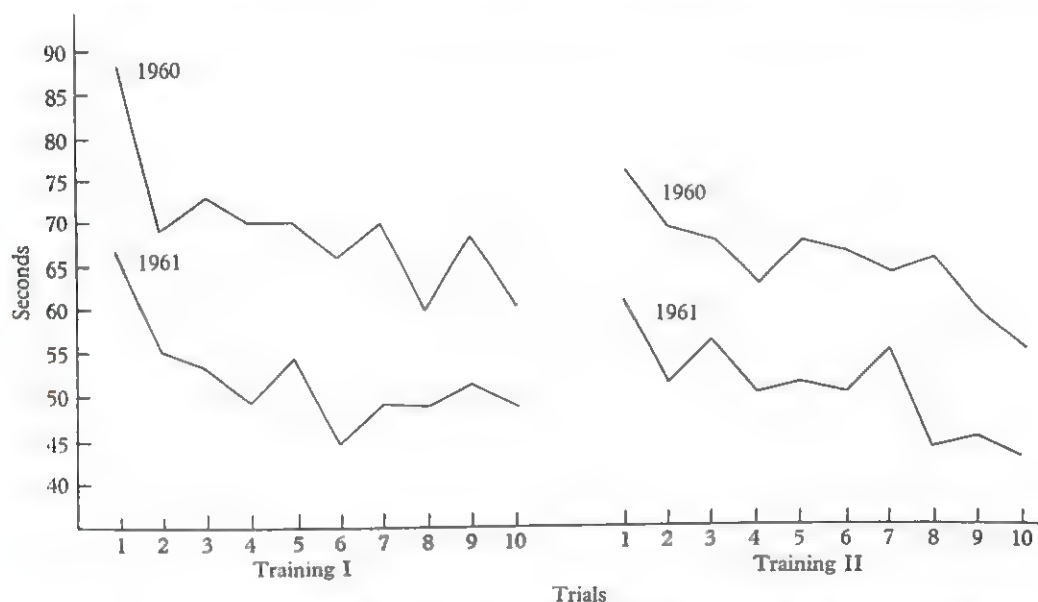


Fig. 5. The sorting of plastic typewriter keys (easier version, Fig. 2 (a)) by 6 adolescent and adult imbeciles in March 1960, and in March 1961, without intervening practice.

noteworthy that in every case the subject commenced the task (which a year earlier had been one of four) without hesitation or comment. Fig. 5 shows the results and it is obvious that the prediction was fulfilled. The difference between the 1960 and 1961 performance is significant at the 0.001 level.

We conclude, once again, that the effects of a few minutes practice per day over a period of 2 weeks carry over a year of non-practice and facilitate relearning. It is considered that the responses of these subjects who had neither the benefit of a pre-test nor instructions confirms remarkably our earlier views on the value of perceptual training, even when given in minimal amounts. Moreover, once again maturation cannot have played any part in these subjects, and of course for the older subjects improvement from a level much nearer the ultimate limits is much more difficult to achieve than from a poor starting point.



## VI. DISCUSSION AND CONCLUSIONS

These four studies have supplemented earlier findings which suggested that perceptual-motor transfer was much more likely in imbecile children than in adults, a hypothesis derived from a study of Hebb's (1949) work. In the first new experiment a task was chosen which maximized the imbeciles' perceptual difficulties, and which, within the administrative time limits, proved too difficult for them and was therefore abandoned. It is proposed, however, to use Itard's technique and later train these subjects by employing an intermediary stage or stages. This may well allow them to succeed with this most difficult material, and in a pilot experiment one child has already learned a previously insoluble task by this method.

The second experiment yielded most unexpected results, showing that a few minutes daily training over a period of a fortnight improved performance on a much more difficult sorting task after a six-month period of non-practice. Retention of learning set and improved perceptual and conceptual processes seemed responsible.

The third experiment, using very difficult formboards, showed precisely the same pattern of transfer indicated in the four experiments reported by Clarke & Blakemore (1961). Children showed massive transfer compared with adults (although in this case adult transfer was by no means negligible) and as a result almost reached adult level.

The fourth experiment showed that the experience of spending a few minutes per day on perceptual-motor tasks a year earlier carried over during this long interval of non-practice and enhanced the subjects' performance.

We have throughout ascribed the transfer effects demonstrated to three main processes: learning set, improved perceptual discrimination and improved conceptual discrimination. For those familiar with Harlow's work (e.g. 1949, 1959) it may have been noticed that we are using 'learning set' in a rather narrower sense than proposed by him. Harlow defines learning set as 'learning how to learn' particular kinds of problems, which 'represents a particular type of transfer of training, the transfer between many problems of a single class...'. Monkeys, for example, can 'learn *any* discrimination problem...with the greatest of ease. Training on several hundred specific problems has not turned the monkey into an automaton exhibiting forced, stereotyped, reflex responses to specific stimuli. These several hundred habits have, instead, made the monkey an adjustable creature with an *increased capacity* to adapt to ever-changing demands of a psychology laboratory environment.'

Now it seems likely that 'learning set' in Harlow's sense could cover all the processes separately discussed in our work. His monkeys, like our imbeciles, clearly built up learned expectancies, as well as a more or less permanent alteration in perceptual and conceptual functioning. But it has seemed desirable to try to subdivide this concept, and we have used the term 'learning set' only to describe the creation of a system of learned expectancies connected with the demands of the particular type of experiment (e.g. sorting). It may be worth giving examples of our evidence for the existence of these processes:

(a) *Learning set.* After 6 month and 1 year intervals of non-practice, respectively, subjects were able to repeat a given type of task without hesitation, and in one experi-

ment, without the repetition of instructions. They came to the task, therefore, knowing what it demanded.

(b) *Improved perceptual discrimination.* The formboard tasks needed the subjects' appreciation of shape, size and orientation, all these processes being initially very poorly developed. Yet there was a carry over of learning in all subjects even when both boards and shapes were changed at the same time (as between Boards II and III). Since learning set appears to be rapidly established (i.e. initial instructions did not need repetition) and since the motor aspects were similar from task to task, it seems probable that improved perceptual discrimination played a part.

(c) *Improved conceptual discrimination.* The earlier paper (1961) gave evidence of the establishment of sorting concepts in the younger subjects. Similarly, in a present experiment the children (like the adults) tended to select several pieces of the same shape, one after another (even though they differed in size) when carrying out the formboards task. Obviously such methods involved concept formation.

The relative importance of these three factors is not known precisely and in any case they are probably linked. Our data suggest, however, that the first two or three trials of any new learning task mainly represent the effect of improvement due to the establishment of set. Since the initial slope of learning curves tends to be steep, this effect is considerable, particularly in unfamiliar tasks. The effect of improved perceptual discrimination seems also to be large, since, as has been implied, subjects continued to improve beyond the initial stages. Probably of least importance are the motor and conceptual elements.

A practical implication of these experiments appears to be that imbeciles benefit markedly from 'sense training'. In fact these and other investigations convince us that the institution imbecile is just about the cleanest of all human slates—he has been deprived of many experiences enjoyed by normal children because of severely damaged cognitive mechanisms which restrict input; he has little initiative which further limits his learning experiences; and above all, in a hospital his environment is also stereotyped and narrow. Hence a given new learned sequence has far wider implications than the same one for a normal. And, as a very slow learner, the imbecile needs more rather than less experience than a normal child to achieve a given goal. In particular these results are completely in line with the views of Itard (often wrongly regarded as an historical curiosity) and hence of Seguin and Montessori. They suggest a deficiency in spontaneous learning from life experiences which can, to some extent, be modified. Among other implications, it seems that the sort of play activities so frequently given to these children in hospital 'schools' can have little effect on their perceptual processes.

Moreover, much of what is commonly called 'formal training' is not training but the imposition of formal discipline on passive subjects. Tizard's recent film (*The Brooklands Experiment*), for example, shows traditional hospital methods of toilet 'training', social 'training' and so forth. Here there appears to be little active participation by the imbecile child, and since he seems not to learn from passive experience, no great progress occurs (see also Tizard, 1960).

Let us now consider more general implications. Previous work has suggested that imbecile learning, in spite of its very impaired range, is rather similar to normal learning except that it occurs in 'slow motion'. Thus the obvious next step is to

repeat these transfer experiments on normal nursery school children. If our results were confirmed, this would lead us to try to find out whether there might be a germ of truth in the theory of formal discipline, not at the later ages at which it was tested, but earlier in childhood. Yet another problem is the validity for normal children of 'sense training', an idea which Montessori took from Itard and which is now somewhat discredited. It is likely that sense training will have little effect on children from intellectually rich environments, simply because they have already had the maximal effects from their surroundings. One would predict, however, that young normal institutional children, or those from adverse homes, would respond and show general transfer. And indeed, according to Kirk (personal communication) the studies of the effects or non-effects on later development of nursery school experience can be interpreted along these lines. These, then, are the sort of questions which these experiments have raised, and which need investigation.

We are very grateful to Dr Ann Clarke for criticisms, detailed suggestions and theoretical contributions; to Mrs P. Higgins for undertaking one of the pilot experiments; to Mr D. Cookson for statistical assistance, and to Dr J. F. MacMahon for constructive comment.

#### REFERENCES

- CLARKE, A. D. B. & BLAKEMORE, C. B. (1961). Age and perceptual-motor transfer in imbeciles. *Brit. J. Psychol.* **52**, 125-31.
- HARLOW, H. F. (1949). The formation of learning sets. *Psychol. Rev.* **56**, 51-65.
- HARLOW, H. F. (1959). Learning set and error factor theory. In Koch (1959).
- HEBB, D. O. (1949). *The Organization of Behaviour*. London: Chapman and Hall.
- ITARD, J. M. G. *The Wild Boy of Aveyron*. Trans. G. and M. Humphrey (1932). New York: Appleton-Century-Crofts.
- KOCH, S. (1959). *Psychology: A Study of a Science*. New York: McGraw-Hill.
- LURIA, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behaviour*. London: Pergamon.
- SHAPIRO, M. B. (1951). Experimental studies of a perceptual anomaly. I. Initial experiments. *J. Ment. Sci.* **97**, 90-110.
- SHAPIRO, M. B. (1952). Experimental studies of a perceptual anomaly. II. Confirmatory and explanatory experiments. *J. Ment. Sci.* **98**, 605-17.
- SHAPIRO, M. B. (1953). Experimental studies of a perceptual anomaly. III. The testing of an explanatory theory. *J. Ment. Sci.* **99**, 394-409.
- TIZARD, J. (1960). Residential care of mentally handicapped children. *Brit. Med. J.* **1**, 1041-6.
- TIZARD, J. & LOOS, F. (1954). The learning of a spatial relations test by adult imbeciles. *Amer. J. Ment. Defic.* **59**, 85-90.

(Manuscript received 12 May 1961)



## CHANGES IN AUTOKINETIC PERCEPTION AS A FUNCTION OF THE TRANSFER OF CONDITIONING EFFECTS\*

BY BOBBY J. FARROW AND JOHN F. SANTOS

*The Menninger Foundation*

This study was designed to test the hypothesis that the spatial region in which autokinetic movement occurs can be altered by means of negative reinforcement operations.

Three groups were employed. Control Group I was tested on autokinetic perception, and after a 4 min. delay was retested. Control Group II viewed and reported the location of a moving light with no negative reinforcement between the test and retest of autokinetic perception. The Experimental Group received the same training as Control Group II, but was negatively reinforced while viewing and reporting the location of a moving light. A reinforcement schedule provided 75% of the reinforcements on the predominant side (left or right) of initial autokinetic movement and 25% of the reinforcements on the opposite side.

The data supported the hypothesis that the spatial region of autokinetic movement can be altered by conditioning operations intervening between test and retest. An explanation of the results in terms of the conditioning of affect to the individual's subjective space is proposed and some implications of these effects for perception and behaviour are discussed.

### I. INTRODUCTION

Over the years autokinesis has proven to be an extremely interesting phenomenon for both clinical and experimental psychology. It has, in fact, a rather long history as psychological techniques go, having been scientifically reported as early as 1799 by Von Humboldt (Adams, 1912). The subjective nature of the phenomenon was first demonstrated by Schweitzer (1858), who found that when several observers viewed a stationary light simultaneously there were discrepancies in the latency and direction of the perceived movement reported. Independent investigators (Voth, 1941; Sexton, 1945) have not only found wide individual differences with this technique but have also obtained very high test-retest reliability for extent and pattern of movement. This subjectivity and reliability presents some fascinating possibilities for its use as a research and diagnostic tool.

The considerable literature on autokinesis also discloses that it may be altered by group influences (Sherif, 1935), that it can be used as a projective technique in which subjects may perceive the motionless light as spelling out words (Rechtschaffen & Mednick, 1955), and that neutral words are seen more easily and frequently than disturbing words (Mednick, Harwood & Wertheim, 1957). Quite naturally, attempts have been made to relate this subjective, flexible experience to personality characteristics and organization. For example, Voth (1947), Sexton (1945), Eysenck, Granger & Brengelmann (1957), Kline (1952) and Gravely (Granger, 1953) have compared indices of apparent movement observed by normal subjects with those of various classifications of psychiatric disorders. These studies generally found differences between patient and normal categories; Gravely (Granger, 1953), however, failed to confirm this finding.

Kline (1952) reported that patients with schizophrenic disorders perceived little

\* This research was supported by Research Grant M-3924 from the National Institute of Mental Health of the National Institutes of Health, United States Public Health Service.

or no movement more frequently than a normal sample. In contrast to these results, Voth (1947), who tested 845 patients, and Sexton (1945), who tested 50 patients, found that those patients with schizoid-type disorders perceived an *exaggerated* amount of movement, whereas those with affective disorders perceived little or no movement. Voth found that as patients 'improved' and approached recovery they perceived more moderate amounts of movement. He also noted that patients who fall in the moderate range initially had a better prognosis for recovery, whether or not they were treated, and that those showing extreme deviations in either direction from the middle range had a poorer prognosis for recovery. Those patients who reported extreme amounts of movement were observed to have the poorest prognosis of all. Voth observed that electro-shock treatment reduced the amount of movement and that it also had the effect of smoothing out the pattern of movement. Sexton (1945) reported that the amount of perceived movement was directly related to the degree of withdrawal from reality and provided an objective measure of introversion-extraversion.

Some recent investigations have suggested that autokinesis may also be a sensitive indicator of certain subjective experiences and emotional states, such as those produced by stress conditions and various isolation procedures. The data indicate that susceptibility to the autokinetic effect is increased following sensory or sleep deprivation (Doane, Mahatoo, Heron & Scott, 1959; Fisher & Rubenstein, 1956). Other studies indicate that when exposure to an autokinetic light source is preceded by combined social and sensory deprivation, the latency in perceiving movement is shorter than when preceded by either social or sensory deprivation alone or by neither of these (Walters & Quinn, 1960).

A good deal of evidence, therefore, suggests that autokinesis is easily produced, is subjective yet reliable, is a sensitive indicator of psychological stress and is a promising technique for personality research (Eysenck *et al.* 1957; Doane *et al.* 1959; Fisher & Rubenstein, 1956; Walters, Marshall & Shooter, 1960). In addition, it has obvious advantages as a technique for demonstrating perceptual learning effects which in turn should have definite implications for behaviour in general and personality characteristics in particular. Most experimental demonstrations of perceptual learning have, in fact, employed relatively ambiguous or unstructured stimuli or experimental situations so that needs or reinforcements may influence the structuring of percepts in such a way as to be easily demonstrable and measurable. With highly structured stimuli, on the other hand, percepts can be influenced relatively little in normal subjects so that one must be content to show that certain aspects are enhanced or strengthened while others are de-emphasized or overlooked.

Although autokinesis has been investigated in a variety of studies, few attempts have been made to demonstrate that specific prior experiences can determine the spatial region within which the apparent movement will occur. The hypothesis to be tested here is that the amount of autokinetic movement perceived in a given spatial region can be substantially decreased by associating negative reinforcement with that region by means of conditioning procedures.

## II. METHOD

*Apparatus*

The apparatus for the training phase consisted of a small electric motor which powered a Graham reduction gear box. A 21 in. arm was attached at its midpoint to the shaft of the gear box. A smaller arm with a pinpoint light source was mounted at the end of the longer arm. The light source was a mercury-cell pen light hooded with black masking tape through which a pin-hole permitted a small amount of light to pass. When the motor was activated, the movement of light at 12 r.p.m. was eccentric, jerky, and roughly circular. A constant voltage d.c. electric stimulator produced a 5 msec. shock which was given by means of palmar electrodes.

The equipment for the test trials consisted of a light source identical with that used on training trials. It was mounted on a stand at a height of 34 in. and was placed at the centre of the circular light movement produced in the training phase.

*Subjects*

Subjects were one male and two female college students and 21 male and 21 female high school students who volunteered. All subjects disclaimed prior knowledge of the autokinetic phenomenon when questioned after the experiment.

*Procedure*

The treatment of the three groups is summarized in Table 1. The test and retest of autokinetic perception were identical for all groups; different conditions were applied only between these two trials.

Table 1. *Treatment by groups showing various phases and conditions*

Group	N	Trial I ( $T_1$ )	Training phase	Trial II ( $T_2$ )
Control Group I	15	Autokinetic perception	4 min. rest; no training involved	Retest of autokinetic perception
Control Group II	15	Autokinetic perception	2 min. of viewing and identifying the location of the moving light source without shock	Retest of autokinetic perception
Experimental Group	15	Autokinetic perception	2 min. of viewing and identifying the location of the moving light source with shock	Retest of autokinetic perception

Each subject was brought into a partially darkened room and seated directly in front of the apparatus (7 ft. from the training light source and 6 ft. 8 in. from the autokinetic light source), which was covered with black cloth when the subject entered. The electrical stimulator remained covered at all times. After the subject had been allowed to relax and adapt to the experimental situation, the room was totally darkened. While dark-adapting, the subject was handed a sheet of paper and these instructions were given:

'In a moment I will show you a light. I want you to watch the light very carefully, and on the paper you just received, trace the movement of the light exactly as you see it move. Do not take your eyes off the light, and follow every movement of the light with your pencil.'

The pencil was then placed in the centre of the sheet and the subject was instructed not to move the pencil until the light moved. The black cloth was removed from the autokinetic light source, which was activated simultaneously with the electric motor (the motor had no function here except to maintain the sound throughout all phases of the experiment and perhaps to suggest movement). After the subject had viewed the stationary light for one minute and recorded its apparent movement, the light and motor were turned off simultaneously and the



test sheet taken from the subject. Subjects were then exposed to one of the following experimental conditions:

(1) *Control Group I*. Fifteen subjects were given a 4 min. time lapse between the two autokinetic trials ( $T_1$  and  $T_2$ ). The time lapse was approximately equal to the total time required for the procedures used with Control Group II and the Experimental Group, and served as a control for the influence of time lapse *per se* on the autokinetic experience. During the delay, the experimenter chatted with the subjects about various topics unrelated to the experiment.

(2) *Control Group II*. Fifteen subjects were given a 2 min. training period between  $T_1$  and  $T_2$  in which they were exposed to circular light movement and were instructed to observe the light carefully and describe its location verbally with reference to the positions on the face of a clock. This group served as a control for the influence of exposure to the moving light on subsequent autokinetic movement.

(3) *Experimental Group*. Fifteen subjects received the same training as those in Control Group II except that they were shocked (no-escape, no-avoidance method) while viewing and reporting the circular light movement. The tolerance level for shock was determined for each subject after  $T_1$  and prior to training. Thresholds for unpleasantness were estimated by slowly increasing the voltage continuously to the right hand through palmar electrodes until the subject reported pain. This was done by having the subject report various levels of the shock: (i) when a tingle sensation was first experienced, (ii) when the shock became unpleasant, and (iii) when the shock became painful. Each subject was then shocked at his subjective threshold for unpleasantness (level (ii)).

A partial reinforcement schedule was employed such that 75 % of the shocks (21) were given on the initially preferred side (as determined from  $T_1$ ) and 25 % of the shocks (7) were given on the non-preferred side. The shocks were concentrated around the three or nine o'clock zones and were given while subjects described the moving light by approximating its successive positions with reference to the face of a clock. Thus, if predominant movement was to the right of centre on the  $T_1$  score sheet, the subject received 21 negative reinforcements to the right (one o'clock to five o'clock inclusive) and 7 negative reinforcements to the left (seven o'clock to eleven o'clock inclusive).

The electrodes were removed before proceeding with  $T_2$  in which each subject was retested for autokinesis using the same procedure as that employed on  $T_1$ .

### III. RESULTS

Separate score sheets were used for  $T_1$  and  $T_2$ . The amount of movement occurring to the right or left of a centre line was measured by means of a map reader. These scores were converted into percentage of total movement for given trials. The score on  $T_1$  involved the percentage of total movement occurring on the preferred side (greater than 50 %). The score on  $T_2$  involved the percentage of movement on the initially preferred side after training or time lapse without training, depending on the particular group involved.

Table 2 shows the percentage of movement occurring in the originally preferred side on both  $T_1$  and  $T_2$  as well as the differences between  $T_1$  and  $T_2$  for all subjects in the three groups. Seven subjects in Control Group I perceived more autokinetic movement to the right on  $T_1$  and 8 perceived more to the left, while for Control Group II there were 9 to the right and 6 to the left; in the Experimental Group there were 6 to the right and 9 to the left. A  $\chi^2$  test was applied to the number of subjects in each group who did not show change of preference and the number of those who did. It was found that 11 Control Group I subjects, 12 Control Group II subjects, and 5 Experimental Group subjects did not change while 4 Control Group I subjects, 3 Control Group II subjects, and 10 Experimental Group subjects did. The analysis yielded a  $\chi^2$  of 10.33 which for 2 D.F. is significant at the 0.01 level. The fact

that 23 Control subjects maintained their original preference on  $T_2$  while only 7 changed, suggests good reliability for autokinetic movement within a given spatial region.

The Kruskal-Wallis test (1952) was applied to the difference scores in Table 2 and yielded an 'H' of 9.97 ( $P < 0.01$ ). Examination of the sums of ranks for the three groups (Control Group I = 274; Control Group II = 285; Experimental Group = 476) revealed that the significance is primarily due to the changes obtained from Experimental subjects.

Table 2. *Individual percentage scores for movement occurring in preferred spatial regions on  $T_1$  and  $T_2$  and the difference scores between these two trials*

Subject	Control Group I			Control Group II			Experimental Group		
	$T_1$	$T_2$	$T_1 - T_2$	$T_1$	$T_2$	$T_1 - T_2$	$T_1$	$T_2$	$T_1 - T_2$
1	95.5	96.8	-1.3	94.1	100.0	-5.9	68.8	52.4	16.4
2	100.0	75.6	24.4	62.5	52.9	9.6	100.0	9.5	90.5
3	100.0	100.0	0.0	53.8	100.0	-46.2	75.0	0.0	75.0
4	100.0	100.0	0.0	100.0	100.0	0.0	63.6	0.0	63.6
5	100.0	79.6	20.4	76.9	50.0	26.9	100.0	0.0	100.0
6	58.3	0.0	58.3	59.5	0.0	59.5	100.0	45.5	54.5
7	69.0	14.3	54.7	66.7	84.2	-17.5	63.6	0.0	63.6
8	64.7	75.0	-9.3	76.2	40.3	35.9	100.0	50.0	50.0
9	100.0	100.0	0.0	100.0	37.5	62.5	57.1	100.0	-42.9
10	83.3	0.0	83.0	100.0	61.1	38.9	100.0	30.8	69.2
11	100.0	100.0	0.0	100.0	51.9	48.1	81.1	56.4	24.7
12	100.0	100.0	0.0	100.0	85.0	15.0	88.1	35.3	52.8
13	100.0	58.8	41.2	100.0	69.0	31.0	100.0	0.0	100.0
14	100.0	46.0	54.0	100.0	100.0	0.0	58.8	0.0	58.8
15	78.9	100.0	-21.1	100.0	60.0	40.0	80.6	52.0	28.6

Table 3. *Mean percentage of movement occurring in the preferred spatial region during  $T_1$  and  $T_2$*

Groups	$T_1$	$T_2$	Difference
Control I	89.98	69.74	20.24
Control II	85.89	66.00	19.89
Experimental	84.06	29.87	54.19

Table 3, which gives the mean percentage of movement occurring on the preferred side on  $T_1$  and  $T_2$  for each group, indicates that the mean percentages for all groups were similar on  $T_1$ , and that Control Group I and Control Group II show similar mean percentages on  $T_2$ . The Experimental Group, however, was markedly different from the two control groups on  $T_2$  (the difference between  $T_1$  and  $T_2$  for the Experimental Group is greater than  $2\frac{1}{2}$  times the difference for either of the control groups).

#### IV. DISCUSSION

The dramatic shifts in the preferred region for autokinetic movement might be explained in a number of ways. One possibility is that through the conditioning operation negative affect was associated with the originally preferred region and thereby altered the individual's subjective space in the autokinetic situation in such

a way as to influence the subsequent autokinetic experience. If, as some experiments indicate (Fisch & McNamara, unpublished manuscript), affect is conditionable to spatial regions, perceived movement might be expected to flow away from an area or region of greater negative affect and toward an area of less negative affect and resistance in a manner similar to escape or avoidance behaviour. Interestingly enough, some of our data indicate that these effects may be temporally labile, in that the autokinetic movement first moved away from the spatial region to which negative affect had been most strongly associated and then returned to the originally preferred side. Perhaps a longer training period and more numerous or intense reinforcements might produce even more striking or lasting effects.

Although the present results were obtained under conditions of stimulus impoverishment, it is possible that they might offer some suggestions for an understanding of certain processes underlying behaviour in more 'natural' circumstances. If affect is conditioned to an individual's subjective space this may contribute to the development of perceptual vigilance or defence, to rigid attention deployment toward or away from certain areas of the perceptual field, and to veridical or non-veridical perceptions of various types and degrees. The nature and extent of the process would depend upon the type and magnitude of affect and the conditions giving rise to it (positive or negative reinforcement). Positive reinforcement should produce alerting effects (Fisch & McNamara, unpublished manuscript) whereas negative reinforcement might give rise to vigilance or defence, etc., depending upon the strength of the reinforcing agent (McNamara, Solley & Long, 1958; Mangan, 1959; Pustell, 1957). At any rate, where the perceptual world of the individual comes to have high affective loading this should have an important influence on the extent and degree to which the environment will be scanned, perceived, attended to, avoided or distorted.

The quantification of perceived autokinetic movement and the meaning of the obtained scores present some very difficult problems which are beyond the scope of the present paper. However, a closely related problem which should be mentioned in passing involves the different subjective scale that may be used in determining and reporting the perceived movement verbally or otherwise. For example, two subjects may perceive identical amounts of movement but one may report this movement with a line of a few inches while the other reports it with a much longer representation. The scaling problem is an interesting question in and of itself and could also have important implications for personality research. This problem was overcome to some extent in the present study by converting movement scores into percentage scores and by using each subject as his own control. Since the intention was to compare movement in opposing spatial regions, the percentage scores equate the amount of movement occurring in a given spatial region for a given subject regardless of the distance of the apparent movement.

In the present experiment an attempt was made to control for the effects of time *per se*, for the effects of viewing and reporting an actually moving light, and for the effects of motor noise on autokinetic movement. An additional control which might be of some importance was not carried out: there is the possibility that the mere determination of shock thresholds in the experimental group might have influenced the pattern, extent or region of movement on the test trials for the experimental subjects. However, it seems unlikely that this factor could account for the differential



changes in region of movement which occurred in the experimental group. A more likely possibility is that extent or pattern of movement would have been influenced, but there was a striking consistency in these for all three groups.

# REFERENCES

- ADAMS, H. F. (1912). Autokinetic sensations. *Psychol. Monogr.* **14**, No. 2 (whole No. 59).
- DOANE, B. K., MAHATOO, W., HERON, W. & SCOTT, J. H. (1959). Changes in perceptual function after isolation. *Canad. J. Psychol.* **13**, 210-19.
- EYSENCK, H. J., GRANGER, G. W. & BRENGELMANN, J. C. (1957). *Maudsley Monographs*, No. 2. New York: Basic Books, Inc.
- FISHER, S. & RUBENSTEIN, I. (1956). Effects of moderate sleep deprivation on social influence in the autokinetic situation. *Walter Reed Army Institute of Research Report*, pp. 63-7.
- GRANGER, G. W. (1953). Personality and visual perception: a review. *J. Ment. Sci.* **99**, 8-43.
- KLINE, N. S. (1952). Vestibular function and autokinesis. In Mettler, F. A. (ed.) *Psychosurgical Problems*. New York: Blakiston.
- KRUSKAL, W. H. & WALLIS, W. H. (1952). Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Assn.* **47**, 583-621.
- MANGAN, G. (1959). The role of punishment in figure-ground reorganization. *J. Exp. Psychol.* **58**, 369-375.
- MCMANARA, H. J., SOLLEY, C. M. & LONG, J. (1958). The effects of punishment (electric shock) on perceptual learning. *J. Abnorm. Soc. Psychol.* **57**, 91-8.
- MEDNICK, S. A., HARWOOD, A. & WERTHEIM, J. (1957). Perception of disturbing and neutral words through the autokinetic word technique. *J. Abnorm. Soc. Psychol.* **55**, 267-9.
- PSTELL, J. E. (1957). The experimental induction of perceptual vigilance and defense. *J. Personality*, **25**, 425-38.
- RECHTSCHAFFEN, A. & MEDNICK, S. A. (1955). The autokinetic word technique. *J. Abnorm. Soc. Psychol.* **51**, 346.
- SCHWEITZER, G. (1858). *Ueber das Sternschwanken*, pp. 477-500. Moscow: Moskovskoe Obshchestvo Ispytatelei Prirody.
- SEXTON, M. C. (1945). The autokinetic test; its value in psychiatric diagnosis and prognosis. *Amer. J. Psychiat.* **102**, 399-402.
- SHERIF, M. (1935). A study of some social factors in perception. *Arch. Psychol.* **27**, 1-60.
- VOTH, A. C. (1941). Individual differences in the autokinetic phenomenon. *J. Exp. Psychol.* **29**, 306-22.
- VOTH, A. C. (1947). An experimental study of mental patients through the autokinetic phenomenon. *Amer. J. Psychiat.* **103**, 793-805.
- WALTERS, R. H., MARSHALL, W. E. & SHOOTER, J. R. (1960). Anxiety, isolation, and susceptibility to social influence. *J. Personality*, **28**, 518-29.
- WALTERS, R. H. & QUINN, M. J. (1960). The effects of social and sensory deprivation on autokinetic judgments. *J. Personality*, **28**, 210-19.

(Manuscript received 5 June 1961)



## PUBLICATIONS RECENTLY RECEIVED

*The Psychology of Expression.* By SYLVIA HONKAVAARA. British Journal of Psychology Monograph Supplements, No. 32. London: Cambridge University Press. 1961. Pp. x + 96. 25s.

The monograph opens with an arresting first sentence: 'This book . . . is a scientific monologue by a heretic', but as the argument unfolds it becomes clear that the thesis is not very heretical, nor is it so scientific as to be unreadable by students of the arts.

The author was provoked to research by some not too well-considered statements of Gestalt psychologists who were impressed by the fact that very young children are quick to perceive human faces and to interpret the expressions on these faces—so much so as to suggest that the understanding of expression is 'innate' and precedes the perception of objective fact. She accordingly embarked on an extensive programme of research (some twenty-odd experiments are reported) designed to establish, and they *do* sufficiently establish, the more orthodox and commonsense view that 'matter of fact' perception comes before the perception of emotions displayed in facial expression. So far, so good. There is no heresy here. But the evidence adduced is stretched to support a larger theory: that there are four phases (or 'dimensions') in the development of perception—the dynamic-affective, the matter-of-fact, the physiognomic, and the intersensory. It is rather hard work to find a quite clear account of these phases or dimensions, and rather hard work to find one's way about the report of the experiments by which the larger thesis is supported. The experiments are divided into 'series', classified according to methods used, and they vary in the age groups and the nationalities of the subjects. The procedure was obviously rather hand to mouth, and conditioned by the available facilities in Finland, England and the United States, where the experiments were carried out. There was clearly no opportunity for a grand experimental design which would make possible an analysis of variance. But Dr Honkavaara cannot be blamed for that. The conditions of research for a lone worker are not conducive to grand experimental designs. In detail the reports of the experiments are extremely fascinating, and are very important for studies both in science and in the arts.

The superstructure of theory is not perhaps sufficiently supported by the evidence contained in this monograph. The four phases (or dimensions) of perception are not defined with sufficient clarity. In fact the discussion of the fourth phase (or dimension) of perception is postponed to a later publication. The review of previous and contemporary research, though fairly comprehensive, overlooks some very important studies. It takes no adequate account of the Lippsian concept of *Einfühlung* or of the concepts involved in projective techniques. There is no reference to Edward Bullough's classical study of perceptual attitudes which must have an important bearing on any theory of the development of perception. It would be interesting to know whether the subjects of Dr Honkavaara's experiments who failed to give the expected responses belong to other than Bullough's objective type, who would have been expected to give 'matter of fact' responses. It is to be hoped that some of these matters will be covered in the promised later report, or better still in a more comprehensive, more systematic work. In the meantime, this monograph can be commended as a quite essential document for all who are engaged in research in the psychology of expression.

C. A. MACE

*Personality Structure and Human Interaction.* By H. GUNTREP. London: Hogarth. 1961. Pp. 456. 45s.

For those who are concerned to develop a more rigorous approach to the study of personality and social interaction, while retaining, or even maximizing, man as a person as the object of study, the clarifications this book offers, and makes possible for the future, are of very great importance.

Dr Guntrip's main aim is 'to trace the way in which psychoanalysis has been in process of outgrowing its origins in a physiological and psychobiological philosophy of man, using the



instinct concept as the basis of theory, into a truly psychodynamic theory of personality implying a philosophy of man that takes account of his reality as an individual person'.

Following introductory chapters in which he outlines some of the theoretical and practical implications of his work with reference to psychoanalysis and its relations with psychology and psychiatry, Dr Guntrip proceeds to examine the beginnings of psychoanalytic theory. Thenceforward he covers in a very closely argued and well-documented critical survey, Freud's early and later theories, examining the contradictions in the basic concepts he used as he 'hovered between a psychology of the organism and a psychology of the person'. After a briefer discussion of the theories of Adler, Horney and Fromm, and Sullivan's Interpersonal Theory of Psychiatry, the author moves on to consider in detail the contributions of Melanie Klein and Fairbairn, wherein, in varying measure, some of the contradictions exposed in the study of Freud's work are resolved in the 'Object Relations' theory of personality. The book ends with a summary discussion of some of the basic forms of human relations, and a final chapter in which some of the theoretical issues are considered in relation to practice in psychoanalysis.

This is a long book. Some readers may feel that Dr Guntrip presents his arguments in too great detail and with too many references and some repetitions. Not only has this method of presentation the merit of reproducing the author's processes of thinking (which are clear-cut and vigorous, and well matched by his lucid and at times incisive style of writing), but it provides also a good learning experience for the reader, for strands of thought are continually unravelled and reknit within different configurations as the main theme of the book develops. The concepts subsumed in the terms ego, id, superego, libido, pleasure principle, reality principle, etc., become very familiar to the reader, as do the contradictions inherent in their older and indeed their current usage. And the arguments are so well developed that such of these terms and concepts as are lost on the way are not at all missed. They are replaced by more satisfying and consistent notions which still stimulate the reader to ask more critical questions, while at the same time providing him with sharper and more suitable tools for doing so.

It is not easy for the serious student, the teacher, or the research worker in general psychiatry, in psychology or social science to become adequately informed about the present status of psychoanalytic theory and the possible direction it may take in future developments. Dr Guntrip's book, perhaps more than any other, makes this task much easier. It is to be hoped that he may be able to provide a somewhat shorter and cheaper edition so that it may be more widely available to students in the social sciences and to the interested reading public.

H. PHILLIPSON

*Conceptual Systems and Personality Organization.* By O. J. HARVEY, D. E. HUNT and H. M. SCHRODER. New York and London: Wiley. 1961. Pp. xiv + 375. 57s.

This book is an attempt to provide what might be called a dialectical theory of stages of development on the way to psychological maturity. The basic idea is the 'concept', by which is meant a kind of schema or 'system of ordering'. The environment is interpreted or responded to in terms of interrelated systems of 'concepts'. The term 'concept' must not be understood in its usual sense, but as indicating a 'mode' or relationship between individual and environment more akin to Kelly's 'construct'. It is perhaps a pity that the term 'concept' should have been chosen for this purpose, since one's familiarity with its more common usage rather tends to confuse one, especially when one reads about the 'confirmation' and 'refutation' of concepts.

The main dimension in terms of which 'concepts' vary is 'concreteness abstractness'. Development is towards the more abstract, and the greatest degree of abstraction is regarded as the most desirable and the ultimate aim of sound therapy and education. There are four main stages. The first is characterized by 'unilateral dependence' or complete dependence on some authority, the second by 'negative independence' which 'represents a lessening of the importance of external control and the initial budding of internal control', the third by 'conditional dependence and mutuality' in which the person 'views other people less subjectively... and more in terms of others' standards and past experience' and the fourth by 'positive independence', which means an 'abstract, interdependent and informational' way of looking at things. In addition, concepts belonging at each stage are thought of as having two 'poles'; progression to the next stage depends upon an 'integration' of both poles. Progression from stages 1 to 2, for example,

involves an integration of 'external control' and 'opposition to external control'. Progression is, furthermore, thought of as proceeding by leaps. Examples are taken from developmental psychology which suggest that the four stages typically correspond to different ages. But there may be 'arrestation' at any stage. This may be either transitional or permanent. If it is permanent, then the person's thinking and attitude continue to show the characteristics of the stage of arrested development. Some account is given of the various aspects of experience and training which encourage progression or arrestation at each stage. The scheme is then applied in the interpretation of personality, normal and abnormal. Gross abnormalities are interpreted as representing extremes of the various types of conceptual functioning.

Although of course the idea of stages of development is far from new, this development of it appears to the reviewer to be an interesting and potentially fruitful one. It is therefore all the more regrettable that the authors have not made a clearer and more convincing exposition of it. The language in which the book is written makes it unnecessarily difficult to read, and in many places is extremely slovenly. Words like 'absolutistically' and 'concretistically' are used without apparent justification; 'concretistically' and 'concretely' are actually used on the same page, seemingly to mean exactly the same thing. 'Alternate' is used when 'alternative' is meant, and so on. The reviewer had to stop and think before coming to the conclusion that in its context the sentence 'Incoming information is not distorted by fitting it onto an absolutistic conceptual schema' meant (to retain the language of the book), 'Incoming information is not fitted onto an absolutistic conceptual schema and is thus not distorted'. If any of his ex-students were to write in this way, the reviewer would feel that he had indeed failed in an essential part of his duties as a teacher. It may be thought that we are making too much of a fuss about this. But this is just the kind of thing which is apt to get psychology a bad name, and there is really no excuse for it. Surely one is entitled to expect any educated person to be able to write clearly, concisely and with a vocabulary no more difficult or outlandish than is absolutely necessary for the subject-matter concerned.

D. GRAHAM

*Personality and Social Interaction.* By R. H. DALTON. Boston: Heath. 1961.  
Pp. x+381. 48s.

How very good the best American psychological writing can be! It seems a pity that the title of this book makes it sound indistinguishable from many others, for the contents are quite distinctive. Not that Prof. Dalton has a new theory to propound: he is an eclectic who draws on many theories to illuminate case studies taken from his varied clinical experience. The distinctiveness lies in his gift of bringing theory to life, by lucid exposition, constructive criticism and concrete application; and in the warm humanity of his writing, whether about patients or about other writers' ideas.

The book is not intended to be a complete account of all theories of personality; nor does it treat of the social interaction of groups. Its main theme is the reciprocal influence of personalities within the family. Three families are presented, through clinical interviews, each over a period of some months at a critical phase in their relationships, and the personality structure and changing behaviour of each member is discussed in the light of concepts drawn from psychoanalysis and other dynamic psychologies, field theory and learning theory. The theoretical problems thus thrown up—such as the nature of interaction between persons, communication, the development of ego and superego, the function of anxiety—are then further considered with reference to the writings of experimental, clinical and philosophical psychologists. In this process large unexplored areas come to light, and some of the ostensibly scientific formulations—even some of those that get through the author's critical sieve—are seen to be pretentiously descriptive rather than explanatory; yet a growing foundation of useful and generally agreed principles emerges. To weave from these many strands a connected, intelligible and readable text, which demonstrates that psychology can really help us to understand human beings, is an achievement for which we may well be grateful.

TERENCE MOORE

*Motivation and Emotion.* By P. T. YOUNG. London and New York: Wiley. 1961.  
Pp. xxiv + 648. 86s.

In keeping with his very broad definition of Motivation—'...all behaviour is motivated, i.e. causally determined' (p. 538)—the author presents us with a textbook of many fields of psychology. Cognitive and perceptual events are said to be regulated by drives and incentives. Learning is necessary for motives to develop; yet learning would itself be under the control of motivation as defined above. This view probably reflects lack of agreement amongst psychologists over definition of the word 'Motivation'.

While subsumed under 'motivation', emotion is defined as '...a strongly visceralized, affective disturbance, originating within the psychological situation, and revealing itself in bodily changes, in behaviour, and in conscious experience' (pp. 597-8). Particular attention is given to research on physiological changes underlying emotional upsets.

With reading suggestions at the end of each of the twelve chapters and a list of questions dealing with the chapter material (pp. 603-8) this book should prove very useful for undergraduate students. University teachers and researchers may be annoyed with the book's unnecessary length, yet will probably welcome its list of close to 1000 references.

H. WALLERSTEIN

*Emotion and Personality.* By MAGDA B. ARNOLD. London: Cassell. 1961. Vol. I, *Psychological Aspects*. Pp. xiv + 296. Vol. II, *Neurological and Physiological Aspects*. Pp. xviii + 430. £5 per set of two volumes.

Although few will agree with everything this remarkable book contains, there is no doubt that it will become a standard work on this subject. The first volume is a critical review of the literature on emotion and a formulation of the author's own views on human thought and emotion. Her theory of the sequence of events which leads from perception to action can be summarized as follows. Perception is followed or completed by a non-reflective judgement, which is associated with feeling or an immediate emotional response. Then follows a recall of the memories of previous events associated with the current percept. This naturally produces an emotional response. This is followed by an action tendency to imagine possible actions and their consequences. In its turn this is followed by an action tendency to an overt action, which is then appraised as good or bad. This appraisal is associated with an emotion and is finally followed by an action which is considered to be good or by the inhibition of an action considered to be bad. This is a reasonable way of arranging the known facts and has some heuristic value.

In the second volume the author postulates the neurophysiological structures which are involved in her sequence of perception, immediate appraisal, reflexion on past events and future action, and behaviour. It is unlikely that many neurophysiologists will agree with her. Nevertheless, her re-interpretations of well-known neurophysiological findings, such as the Klüver-Bucy syndrome, are very interesting. It is to be hoped that specialists in this field will consider her arguments very carefully. Many interesting side issues arise in this discussion of the physical basis of emotion. Thus the author's suggested physiological basis of memory disarms much of the criticism of the concept of engrams. However, the complete rejection of Magoun's non-specific activating system is hardly justified. The author's theories do not explain the more recent work of Lindley and his group, who have shown that stimulation of the mid-brain reticular system increases the accuracy of visual perception.

There are a few minor points of inaccuracy, such as an incorrect statement of Hobb's concept of phase sequence. However, when the work is taken as a whole, Prof. Arnold is to be congratulated on her masterly survey of the psychological and physiological aspects of emotion. Any psychologist, psychiatrist or neurophysiologist, who reads this book, cannot fail to be stimulated by the many interesting ideas put forward.

F. J. FISH



*Psychological Testing*. By ANNE ANASTASI. Second edition. New York: The Macmillan Company. 1961. Pp. xiv + 657. 56s.

*Measurement and Evaluation in Psychology and Education*. By ROBERT L. THORNDIKE and ELIZABETH HAGEN. Second edition. New York and London: Wiley, 1961. Pp. viii + 602. 58s.

*Tests in Print: A comprehensive Bibliography of Tests for use in Education, Psychology and Industry*. Edited by OSCAR K. BUROS. Highland Park, New Jersey: The Gryphon Press. 1961. Pp. xxx + 479. \$7.00.

New editions of standard textbooks, particularly those used largely for reference purposes, are a mixed blessing. One welcomes the opportunity of being brought up to date, but often the changes—in spite of claims to the contrary—are slight, and sometimes, as far as arrangement goes, even arbitrary. Neither of the above second editions can be entirely exempt from these strictures, but at the same time each is probably fully justified; as Prof. Anastasi says, 'the pace at which psychological testing is developing can be gauged from the fact that about a third of the tests discussed in this edition have either originated or been revised since the publication of the first edition'.

Sections of both books have also been rewritten, in the case of the Thorndike and Hagen book perhaps more noticeably so, but it is remarkable how closely parallel, in general, the changes are. The scope and content of the two books are of course also very similar: both deal with the principles of psychological testing and survey the available tests. The Thorndike and Hagen book is perhaps the simpler, with more emphasis on how to handle psychological test data, and on the devising of objective 'tests' (i.e. multiple-choice examinations) for classroom use. Anastasi's is the more critical: the author's evaluations of test procedures are sound and as a rule very fair and unbiased; occasionally she hits hard, witness the following summing-up of a well-known projective technique (new wording in italics; from the first edition in brackets):

'The X test is *probably* (appears to be) one of the least *promising* (acceptable) of the currently popular projective techniques. Its theoretical *rationale* (orientation)... *appears* (is) particularly weak and *farfetched* (fallacious). Attempts at empirical validation of its various assumptions have so far yielded *overwhelmingly* (consistently) negative results'.

Both books fulfil their aims admirably—Anastasi's is particularly comprehensive—but both have the disadvantage, in this country, that tests of other than American origin are practically ignored. Thus, there is no mention of such well-known British tests, to take only a few examples from contrasted fields, as the Object Relations Technique, the Maudsley Personality Inventory or the Mill Hill Vocabulary Scale. (Incidentally, the word 'Vocabulary' does not appear in either Index.)

No such reproach can be laid at the door of *Tests in Print*. All three tests are there, together with every other commercially available test which the reviewer has looked up. The principal aim of *Tests in Print* is implicit in its title—to present a comprehensive guide to all tests currently available. Of these there are over 2000 listed. Very wisely it was decided to include a section covering tests Out of Print; these number over 800.

No evaluation of individual tests is attempted, but since the book serves also as a master index to the *Mental Measurement Yearbooks*, this is no loss, and the resulting volume is compact and easily handled. Also included, and serving a similar end, is the valuable statement on 'Technical Recommendations for Psychological Tests and Diagnostic Techniques', prepared by the APA and other bodies and originally published in the *Psychological Bulletin*.

The meticulous accuracy that characterizes the *Yearbooks* is equally in evidence here.

B. SEMEONOFF

*Speech Disorders: Aphasia, Apraxia and Agnosia*. By SIR RUSSELL BRAIN. London: Butterworths. 1961. Pp. viii + 184. 42s.

Sir Russell Brain has written an admirable short account of aphasia and related disorders which should be of great value to psychologists, clinical and otherwise. After a brief considera-

tion of the origin and development of speech, Sir Russell deals successively with handedness and cerebral dominance, the history of thought about aphasia, the neurology of language, and the clinical manifestations, prognosis and course of aphasic disorders. Developmental disorders of language (which have important educational implications) receive brief mention. Finally, two chapters are devoted to the agnosias and apraxias, very little known to psychologists but of real importance for our understanding of perception and voluntary action. Although the treatment is at times so compressed as to verge on the dogmatic, there is no dogmatism about the author's theoretical outlook. Indeed his fairness and open-mindedness towards the warring theorists of aphasia, past and present, are throughout above reproach.

In his general approach to the neurology of language, Sir Russell leans heavily on the concept of the 'schema', originally put forward by Head and Holmes to account for the appreciation of posture and passive movement. His interpretation of the concept, however, is much closer to that of Wolters than of its original sponsors, and it is a pity that no reference is made to the two papers in this *Journal* (particularly the second) in which Wolters developed his ideas. Sir Russell, though, goes well beyond Wolters in viewing the 'schema' in terms of probability theory (here he is somewhat reminiscent of Oldfield) and clearly attaches great value to communications theory as a tool in the analysis of speech and its disorders.

It is noteworthy that Sir Russell does not consider that the study of aphasia (or related manifestation of brain disorder) is likely to throw light on the organization of psychological function in the healthy individual. Rather, he suggests, is it the other way round: as our knowledge of language—and in particular its physiological basis—advances, so may we hope to explain the phenomena of disordered speech. Although this view has much to commend it, it should be borne in mind that pathological study may clarify, even if it does not explain, normal psychological activity. One may recall, for example, Hughlings Jackson's celebrated distinction between 'propositional' and 'emotional' speech, Ribot's 'law of regression' in the dissolution of memory, and Head's emphasis on the symbolic mediation of certain types of non-verbal performance. None of these principles, which have undoubted application to healthy psychological function, could have been formulated had inquiry been restricted to those with intact brains. At the same time, Sir Russell is almost certainly right in warning us that clinical study in and by itself is no Royal Road to psychological truth.

It is to be hoped that this book will be widely read by psychologists, and not merely by those whose professional activities bring them into contact with problems of aphasia. The gap between neurology and psychology seems, if anything, to have widened in recent years and this would hardly appear in the best interests of either discipline. For instance, few students of psychology today seem aware of the highly circumscribed defects which may follow localized brain injury or to have considered the implications of such defects for theories of intelligence. Equally, few neurologists (the author of this book is an honourable exception) are at home with recent developments of method and theory in experimental psychology. Yet there is good will on both sides and this book may perhaps help to bring about a *rapprochement*.

Sir Russell Brain represents the high intellectual tradition in British neurology which has contributed so much not only to scientific medicine but also to the foundations of psychology in this country. As we have recently been reminded, Hughlings Jackson was the first Honorary Member of the British Psychological Society and his work, especially through its influence on Henry Head, contributed in no uncertain way to modern psychological thought. Rivers, too, owed much of his outlook to his early training in neurology. Is it too much to hope that psychology, despite its current preoccupation with what Hebb has called the 'conceptual nervous system', may still draw inspiration from neurological reality?

O. L. ZANGWILL

*Research on the Etiology of Schizophrenia.* By G. YU. MALIS. Translated by B. Haigh. New York: Consultants Bureau Enterprises. 1961. Pp. xii + 195. \$9.50.

This volume, first published in Russia in 1959, is a detailed report on the investigations into the etiology of schizophrenia carried out by the late Prof. Malis at the Bekhterev Brain Institute during the period 1938 to 1957. The author's main thesis, that schizophrenia is an infectious process produced by a virus, is a distinct and rather startling departure from the now generally accepted view of an endogenous basis to the illness. Prof. Malis reached his conclusions after detailed studies of the effect of schizophrenic blood serum on the development of tadpoles and plants.

The author's discussion of his own investigations is of a highly technical nature and could be properly evaluated only by biologists experienced in this particular field. It is apparent, however, that, in his discussion and subsequent dismissal of the studies which have suggested an endogenous metabolic factor, the author fails to consider much that is relevant in reports of Western literature during the last decade. Perhaps the most obvious omission is Prof. Malis's lack of reference to any of the many studies on adrenal physiology or to the work on ataraxic drugs, which has been extensively reported by American and European workers. The role of genetic factors in schizophrenia receives very cursory consideration and the findings of such authorities as Kallman are not mentioned. Although some of Prof. Malis's earlier experiments have been reproduced with negative results by American biologists, these studies are not reported in the text.

These apparent limitations only serve to illustrate the deplorable lack of communication between Russian and Western scientists and it is quite probable that many of the interesting Russian biological studies of schizophrenia reported by the author are in their turn unknown to American and European workers. In a short preface to the book, Dr Hoagland, Director of the Worcester Foundation for Experimental Biology, comments that one's disagreement with Prof. Malis's interpretation of his findings cannot affect the importance of his data. Many of the clinical observations reported in the book (e.g. the well-documented evidence of a febrile state preceding the onset of schizophrenia) suggest fruitful fields for further observation and research.

A. MCGHIE

*A Model of the Mind Explored by Hypnotically Controlled Experiments and Examined for its Psychodynamic Implications.* By G. S. BLUM. London and New York: Wiley. 1961. Pp. xii + 229. 60s.

Any book with such a title, in which experiments with hypnotized subjects are described, must start under a handicap. I did my best to read it conscientiously, but became more and more discouraged.

There are diagrams in which arrows join circles or boxes labelled with phrases such as 'Cognitive Representation Anxiety', 'Affective Circuit Anxiety' or 'Primitive thoughts about smashing sib' inside a 'Pathogenic Cognitive Network'. A major part of the text consists of verbatim transcripts, with lavish use of block capitals for emphasis, of dialogues in brisk campus idiom between hypnotist and 'deeply sleeping' subjects (if only people would not mix up sleep with hypnotism). I find it impossible to summarize the theme of the book. The author made no discernible attempt to do so—'Where should we saw off segments of the Mental System in order to have them serve as efficient building blocks?'

The prose is exemplified in the sentence, 'Similarly, the phenomenon of *isolation* of emotion pertains to a sequential process whereby inhibition is invoked after the thought signal is hyper-facilitated and prior to affective discharge'. Obscured within the book there may be, as the author believes, the seeds of a future revolution in psychology. In the absence of discipline, clarity and economy in writing, they remain hidden.

IAN OSWALD

*Handbook of Abnormal Psychology. An Experimental Approach.* Edited by H. J. EYSENCK. London: Pitman. 1960. Pp. xvi + 816. £6.

There has long been a need for an inclusive text of abnormal psychology which could encompass, within the covers of one volume, all the more recent and more important experimental work and which could function as a basic textbook for both students and clinicians in this rapidly growing branch of psychology. With the appearance of this book of nearly 800 closely printed double column pages it may well be thought that this long-felt need is at last satisfied.

In his introduction the editor sets forth his aims in publishing the *Handbook*. These may be summarized in his own words—'Accumulation of facts, while indispensable, is not enough in science... hence we have tried to use the laws and hypotheses of general experimental psychology to account for the facts presented... there is throughout this book a stress on "integration through theory", an attempt to see abnormal psychology as a part of general, experimental



psychology'. The task of reviewing this book would seem to involve an estimate of the extent to which the editor has achieved his stated goal and also an estimate of the extent to which the *Handbook* meets the needs of those engaged in teaching, research and clinical practice in abnormal psychology.

The *Handbook* contains twenty chapters, contributed by nineteen different authors. It is divided into three parts, Part 1 being concerned with description and measurement, Part 2 with causes and determinants and Part 3 with the experimental study and modification of abnormal behaviour. There is an ample bibliography at the end of each chapter, and although some deal with closely related areas of study, collaboration between the different authors has prevented any undue overlap.

Although generally in this book there is a combination of both fact and theory, some parts are more obviously concerned with theory, whereas others, almost entirely factual, are in the form of brief summaries of research and experiment in both general and abnormal psychology. This is done largely without bias in selection and is reported concisely but with enough relevant detail. There is never any difficulty in finding the full reference. Although these chapters may frequently contain material which has been published elsewhere, it is here brought together and in some cases revised and considerably expanded. Two such chapters, 'Abnormalities of Psychomotor Functions', by Yates, and 'Cognitive Abnormalities', by Payne, are to be found in Part 1.

Among the chapters devoted more to theory are those contributed by Prof. Eysenck himself. His opening chapter on 'Classification and the Problem of Diagnosis' criticizes present-day psychiatric nosology and reasons that the neuroses and psychoses are only classed as diseases through historical accident. In place of current psychiatric classifications a system derived from canonical variable analysis is substituted. In a later chapter (in Part 3) the same author discusses 'The Effects of Psychotherapy', with reference to the comparative recovery rates between treated and untreated neurotics. Much of this material has been published elsewhere in both article and book form, and those familiar with Prof. Eysenck's work will find little that is new, either in the argument or in its conclusion.

Another chapter written almost entirely from a theoretical point of view is that contributed by Furneaux. Here there is a discussion of the difficulties involved in the measurement of 'Intellectual Abilities and Problem Solving Behaviour,' and a theoretical model, the 'human problem-solver', is set up. This is done with some humility and the author gives full recognition to the over-simplification involved. The usefulness of the model is confined to 'highly structured tasks within which opportunities for differences in individual approach are as completely eliminated as may be possible'. Within these limits the chapter is a stimulating one. At the same time it focuses attention on what might be regarded by some as a basic flaw in the book as a whole. There is nowhere any serious indication that as well as the problem solver the structure of the problem may be amenable to scientific study. In view of the limitations which Furneaux sets on the 'human problem-solver', the failure to consider the structure of the problem must strictly limit the behaviours which can be investigated and preclude any possibility of attempting to understand human behaviour as opposed to the behaviour of the 'human problem-solver'.

These chapters, which deal primarily with the *Handbook's* theoretical position, show that the editor's stated aim of integration through theory is carried out at one level only rather than at several. It may be doubted whether contemporary psychology can at present claim one all-inclusive theoretical system and many readers might have wished to see a broader theoretical approach and even perhaps one slightly less rigid but more directly relevant to the clinical problems of the individual patient. If one accepts Prof. Eysenck's approach in its entirety, nothing in this book will offend, and the editor will undoubtedly be thought to have achieved his aim. If one does not, then it might be said that this determined adherence to one theoretical position in the *Handbook* is its greatest weakness.

Part 2 includes an excellent discussion of recent research and development in the chapter on 'Heredity and Psychological Abnormality'. This is an area in which our knowledge still has many gaps and one may or may not agree with the author's conclusion that as the geneticist's knowledge of genes and how they manifest themselves increases, there will be correspondingly less significance in the psychologist's attempts to understand the relative contribution of environment and heredity. O'Connor and Franks are co-authors of a chapter on 'Childhood, Upbringing and other Environmental Factors'. This obviously could not be a comprehensive account and instead 'several topics have been selected either because they cover subjects which

at present dominate the literature on children's mental ill-health or because they represent new contributions...to this field'. Within these limits there is a very satisfactory and thorough treatment of the subjects chosen.

Again, in Part 2, a chapter on 'Somatic Reactivity' by Martin provides a brief, although useful, synthesis of research in this clinically important area. 'Psychological Effects of Brain Damage' are discussed by V. Meyer. The chapter includes sections on theories of brain function and organization; claims for localization of function and a critique. The information in this chapter is fairly complete and well ordered and clearly maps out some of the main difficulties in trying to relate psychological deficit to neurological lesion. Such pitfalls are all too often not sufficiently recognized in psychological studies of brain-damaged subjects. The description of 'The Psychological Effects of Cerebral Electroshock' collates data from many different sources and results in an interesting and informative contribution in an area where our present knowledge is again insufficient. A chapter on 'The Effect of Drugs on Behaviour', by Tranton and Eysenck, is also comprehensive and of undoubted value to psychologists who attempt clinical assessments of drug effects in a mental hospital setting.

The *Handbook* is a large and impressive volume; it is also a very expensive one. It is never an easy matter to set the scales and weigh the balance of merit against cost. This is particularly true when, as in this book, the extremes of both variables tend to be reached. As far as merit is concerned, we may call attention to Prof. Eysenck's statement in his introduction that 'psychology in its present humble state can at best support on a factual basis certain low order generalizations' one may feel that many generalizations rather than few, even if they are conflicting, are more likely to promote and encourage advance in psychology along a broad front. New data should not only support old hypotheses but should, where necessary, invoke new ones. On the other hand it cannot be denied that the book presents a remarkable accumulation of facts. If for no other merit than this the *Handbook* has an established place in the literature and is likely to prove invaluable as work of reference for a long time to come. As far as cost is concerned the *Handbook* is certainly expensive, but those who are engaged in teaching or research in abnormal psychology may well feel that they cannot afford to be without it.

E. L. R. MACPHERSON

*An Approach to Cybernetics.* By GORDON PASK. London: Hutchinson. 1961. Pp. 128. 12s. 6d.

Applied mathematics cuts across the usual classification of the sciences. It is concerned with laws and rules that characterize mechanical structures, the flow of fluids, electrical functioning and so forth. Cybernetics, too, is said to cut across the traditional classifications. It, too, is said to be concerned with certain abstract principles that enter into engineering, biology, economics and other disciplines. These principles specify stability and control processes, and perhaps also all forms of communication. However, the subject-matter of cybernetics is as yet ill-defined.

There is probably not a chapter in this book that could not have been dealt with quite satisfactorily within the framework of one or another existing discipline, such as statistics, electrical engineering, physiology or psychology. It is difficult to see what cyberneticians can achieve as such rather than as engineers, biologists, etc., when dealing with, for instance, stochastic models, or biological controllers, or teaching machines, or the structure of industry. Nevertheless, it is laudable that there are some imaginative and enterprising scholars who regard it as their main task to seek 'unifying principles', or abstract 'a controllable system', or otherwise attempt to attain the special aims of cybernetics, however these may be described.

The author devotes the bulk of the monograph to a clear exposition of information theory, theories of 'state determined behaviour', and the principles of control systems. He provides a helpful introduction concerning the origins of cybernetics. The last short chapter on industrial cybernetics is uncomfortably general and a little flamboyant. There is a short glossary of technical terms, a list of references, and several useful appendices. Anyone wishing to know how cybernetics is progressing should read this book.

W. SLUCKIN

*The Nervous System.* By G. M. WYBURN. London and New York: Academic Press. 1960. Pp. viii + 184. 35s.

This book, by the Professor of Anatomy at the University of Glasgow, is designed to give a combined introduction to the structure and function of the nervous system for 'the wide range of non-medical auxiliaries now being recruited into the National Health Service... Physiotherapists, Speech Therapists, Occupational Therapists, and Psychotherapists, to which can be added Clinical Psychologists, Dieticians, Chiropodists, Radiographers, Audiometricians, and Orthopticians', and also 'Food Scientists, Pharmacists, Pharmacologists, Bachelors of Education, and graduating degrees combining training in nursing and physical education', as well as being 'a helpful introduction to the medical student beginning the study of a difficult subject'; the author's experience with 'medical graduates studying for higher diplomas in psychological medicine suggests it would be a useful starting point for many of them'.

The book, therefore, casts its net widely; and the reviewer could not claim to be able to judge its suitability for all these purposes. There is certainly need for a good non-technical primer in this field, which is not merely an abbreviated version of orthodox descriptive anatomy. Among its excellent features are its brief but clear introduction to basic neurophysiology, a useful account of sensory mechanisms, necessarily somewhat dogmatic, and a short but intelligible chapter on the electroencephalogram and the 'reticular formation'. However, the determination to exclude unnecessary detail could well have been extended to the diagrams, some of which seem to have been carelessly copied. It is surely misleading to refer in text and plates to the posterior limb of the lateral sulcus (alias 'Sylvian'), as the 'transverse fissure'; and to depict the major contribution to the 'basilar nerve' as the anterior spinal artery, leaving the vertebrals unlabelled, even though these are not matters of any moment for most of the types of student envisaged, and readily corrected by the rest.

For the first-year undergraduate student of psychology this book does not appear to have any major advantage over the corresponding chapters in his standard textbooks. For the later years the texts of physiological and experimental psychology provide enough at the simpler levels. Honours and clinical psychology students would, surely, prefer to refer to the standard larger works. Students of speech therapy, one supposes, would require more than six pages on speech mechanisms and speech disorders, although grateful to have the rest expertly summarized. However, the author's own teaching experience must have enabled him to adapt his material to the other types of student listed, who might find it difficult to find or use other sources of information. The reference lists at the ends of the chapters will enable the interested student to extend his range, and are not so long as to put him off the project.

J. A. L. NAUGHTON

*From Adolescent to Adult.* By P. M. SYMONDS with A. R. JENSEN. New York: Columbia University Press. London: Oxford University Press. 1961. Pp. xii + 413. 70s.

The late P. M. Symonds belonged to that band of admirable American psychologists who, in the course of the last twenty or thirty years, have endeavoured to combine the merits of an objective and experimental approach to psychology with an interest in the less calculable emotional and social factors and an open mind about Freudian concepts. Stemming from this undogmatic outlook two valuable sets of studies, each now represented by two books, are Symonds's most lasting contributions to the literature of psychology. The first two (*The Psychology of Parent-Child Relationships* and *The Dynamics of Parent-Child Relationships*) are among the few systematic and comprehensive investigations on parent-child relationships. The second set consists of two works on fantasy in adolescent and adult life: one of these was published in 1949; the work under review is a follow-up study of this.

The earlier book, *Adolescent Fantasy* was based on a detailed study of forty boys and girls. The field work for it which had been carried in 1940-41 introduced, in addition to interviews and an interest inventory, the Picture-Story Test, a variant of the thematic apperception method. In 1953 twenty-eight subjects of the earlier research were located. The Picture-Story Test was repeated, each subject was interviewed again and, in addition, given a Rorschach test. Most of



the field work and the analysis of the data were undertaken by the co-author, A. R. Jensen of the University of California.

The result of this follow-up is an intelligent and civilized book which will be read with interest by many psychologists and students. Its merits are several. Above all it contributes to an understanding of the growth from adolescence to adulthood. Here the concept of adolescence as a distinct period, completely different from other stages of life, recedes. The persistence of individual characteristics, the stability of fantasy, the living-out of fantasy in life experience are strikingly observed. Moreover, the interpretation of projective material, in particular of the Picture-Story Test, can be studied in this book. The work further offers an opportunity for stimulating comparisons between earlier and later projections, between Rorschach and the Picture-Story Test and between projective techniques and life experience. It also discusses the prediction from fantasy material of personality and mental health, and the characteristics and significance of fantasy. Welcome are, finally, the *résumés* on each subject and the detailed account of one case in the appendixes. The absence even in outline or reduced form of the pictures which play such an important part in the study is a disadvantage. Apart from that the work can be entirely recommended as a sensitive, searching and balanced contribution to the psychology of development and personality.

H. H. STERN

*Family Environment and Delinquency.* By SHELDON and ELEANOR GLUECK. London: Routledge and Kegan Paul, 1962. Pp. xii + 328. 35s.

Using their data on 500 delinquents and controls the Gluecks set out to apportion constitutional (hereditary) or environmental origins to 67 'traits'. Apart from the methodological objections that can be made against rating by traits, those used by the Gluecks are heterogeneous characteristics devoid of logical principle and such as could be multiplied endlessly. Examples are, Poor Health in Infancy, Banality, Marked Power of Observation, Marked Feeling of Not Being Taken Seriously, Intuition, Unmethodical Approach to Problems.

Two criteria of constitutional or hereditary origin are used. The first is association with one of the four somatotypes, it being assumed that this indicates genetic linkage. The second is independence of the forty-four environmental factors. If these criteria were valid they should corroborate each other. But of the twenty-three somatotype-associated traits only seven (30.4%) meet the Gluecks' criterion of being associated with no or not more than one environmental factor. This compares with 30.3% for the traits independent of somatotype. The Gluecks elude this awkward even chance by discounting four of the somatotype-associated traits as not being causal.

The degree of association of each of the sixty-seven traits with each of the forty-four environmental factors among the delinquents is given in 35 pages of tables showing 258 associations significant up to the 10% level. This may impress the reader who does not realize that with 2948 possible commutations mere chance would give 295 as 'significant' by the modest criterion they use. At the 1% level there were only seventeen as against an expectation of twenty-nine. The Gluecks recommend these results as a basis for diagnosis and treatment.

Apart from this grandiose mishandling of statistical methods, it would be hard to find a book containing so much repetition, and unabashed self-advertisement by its authors.

D. H. STOTT

*Ageing and the Semi-skilled: a Survey in Manufacturing Industry on Merseyside.* (Medical Research Council Memorandum, No. 40.) By ALASTAIR HERON and SHEILA M. CHOWN, assisted by M. S. FEATHERSTONE and CATHERINE M. CUNNINGHAM. London: H. M. Stationery Office, 1961. Pp. viii + 59. 5s.

This is mainly an account of a survey of opinions given by between 200 and 300 departmental managers and foremen in twenty factories on Merseyside about the performance and employment of older semi-skilled men. It also reports the results of a study in which the ages of operatives were related to assessments of the demands made by 526 jobs supervised by the foremen interviewed, and of the environmental conditions under which these jobs were carried out.

It must be admitted that the opinion survey is disappointing and, with the wisdom of hindsight, one may question whether this type of approach is really suitable for studying problems of the employment of older people. To quote just one example, the statement on p. 10 'About 40 per cent (of the managers and foremen) thought that older men had fewer accidents', appears pitifully weak when compared with the several distinguished statistical studies of accidents in relation to age which have been published during the last 40 or so years.

The job study is more valuable, providing much-needed, even if somewhat tentative, additions to existing field evidence about the effects of certain job demands on older people, and suggesting some interesting parallels between field and laboratory findings.

A. T. WELFORD

*Letters of Sigmund Freud, 1873-1939.* London: Hogarth Press, 1961. Pp. 464. 50s.

It can safely be said that Freud as a young man had never heard of psychoanalysis, nor had he been even indirectly under its influence. His intimate letters to his fiancée Martha Bernays (the first hundred or so of this collection) give us who value Freud for his contribution to science and to therapeutics a new chance, better even than the Jones's biography has given us, to judge Freud as a man. Does he turn out to be a human human being? These letters provide an answer.

Freud was born in 1856 and he became engaged to the 21-year-old Martha at the age of 26 in 1882. We are lucky to know so much about this young doctor when he was in love. As a man in love he shows all the usual signs, so that anyone who has been similarly affected must quickly recognize the affection.

We find he has a double allegiance, and he thus addresses himself to 'lofty science': 'Your Highness, I remain your humble, most devoted servant, but please don't hold it against me; you have never looked kindly upon me, never said a comforting word to me; you don't answer when I write to you, listen when I speak, but I know another lady to whom I mean more than I do to you, who repays my every service a hundredfold, and who moreover has but one servant and not, like you, thousands. You will understand if I now devote myself to the other so undemanding and gracious lady. Keep me in pleasant memory until I return. I have to write to Martha' (p. 29). Martha seems to have accepted this fact of the double allegiance. At a later date Freud refers to Martha's own ambition, and it is to be supposed that this found full scope in Freud's ultimate success. But one can see how easily Freud's scientific work could have been destroyed had Martha needed more immediate satisfaction for her ambitions.

In those early days Freud collected round him a small social group composed of Martha's brother and her sister's fiancé and some of his own friends. This was called the Bund, and one wonders whether it might be possible to relate this social group with the later Bund, the professional one, which was composed of certain colleagues who helped to form a psychoanalytical society.

These letters show how Freud, very much a Jew, had a need and perhaps a tendency to find, day by day, the milieu in which he could have freedom, freedom to be himself and freedom to develop in his own way. So he was remarkably independent as a grown person perhaps because of his own dependence on a group, a group that was ever-changing, a group that ultimately depended in its turn at any one moment on Freud's own choice, and on his tolerance and on his intolerance.

But Freud had hidden away in him a terrific potential. It is true that he wrote in obvious sincerity: 'I consider it a great misfortune that Nature has not granted me that indefinite something which attracts people. I believe it is this lack more than any other which has deprived me of a rosy existence. It has taken me so long to win my friends. I have had to struggle so long for my precious girl, and each time I meet someone I realise that an impulse, which defies analysis, leads that person to under-estimate me' (p. 211). Also, of himself '... who is still young and yet has never felt young' (p. 138). And again: '... in my youth I was never young and now that I am entering the age of maturity I cannot mature properly' (p. 214). But he was glad that Breuer said: '... hidden under the surface of timidity there lay in me an extremely daring and fearless human being. I had always thought so, but never dared tell anyone. I have often felt as though I had inherited all the defiance and all the passions with which our ancestors defended their Temple and could gladly sacrifice my life for one great moment in history' (p. 215). Added to this he could make a definite statement like that contained in his letter to Barbara Low after David Eder's death, 'We were both Jews and knew of each other that we

carried that miraculous thing in common, which—inaccessible to any analysis so far—makes the Jew' (p. 424). (By the way, he did in the end try to get behind this very thing when he put forward the thesis that Moses the Egyptian invented the Jew.)

So in this way we get from the letters the picture of a man establishing the personal fact: *I am*; and surely this is the original reason for the anxieties that are loosely called paranoid, and that lead to the necessity for a man to arrange for a circle of friends, both as a defence from all the others who are playing at the life-game 'I'm the king of the castle', and in a positive way as a group of persons safe to love. This seems to be the task of the male half in Plato's fabulous division, and to correspond to the complementary task of the female half, which is to form a circle round the newly conceived.

It was into this pattern that Martha evidently fitted well. She retained the position of being Freud's other half. Freud was by nature the very opposite of promiscuous. Perhaps this one fact more than any other gave Freud the right to startle and disturb the world's mind, and made possible the launching of such ships as *The Dynamic Unconscious*, *Infantile Sexuality*, *The Oedipus Complex*, *Psychic Reality*, and many others that have proved to be not so much pleasure steamers as battleships in the Truth war.

For Freud surely this happy union with Martha to some extent solved his personal problem of bisexuality. He referred to Martha as Cordelia, and in so doing he found himself very happily in the same boat with Breuer (pp. 55, 56). But surely this leaves it for us to work out the place of Goneril and Regan. Lear had to suffer the machinations of his ambisexual first and second-born before he could reach to the essentially female Cordelia. 'Her voice was ever soft, gentle and low, an excellent thing in woman.'

If Freud achieved this simplification in his life by the legitimate means of a happy marriage he did also give us the instrument, the psychoanalytic technique, by which we might go further into the matter of the fear of WOMAN, as she can be in the unconscious of any person, man or woman.

Freud's last letters to Martha before their marriage are evidently missing. The reader suddenly finds Freud writing to his mother-in-law reporting the arrival of a baby. He has suffered during the birth, as men do, and now he is already very much loving his baby girl who is only five hours old. He reports like a good pupil in infant-observation: 'She weighs nearly 7 lb., which is quite respectable, looks terribly ugly, has been sucking at her right hand from the first moment, seems otherwise to be very good-tempered and behaves as though she really feels at home here. In spite of her splendid voice she doesn't cry much, looks very happy, lies snugly in her magnificent pram and doesn't give any impression of being upset by her great adventure' (p. 233).

All this feels very natural and normal, and forms a proper background for the other letters that gradually lead into the territory that has already been explored by Jones.

This reviewer finds immense wealth in this book, and in the insight it gives into the personality of so remarkable a man. Freud could talk about giving 'free play to all the fountains of my irresponsibility' (p. 175); he could write: '...at heart I am still a child; I can be so happy simply because I am in another place, have different money in my pocket' (p. 92). He could laugh a whole evening at a Rome performance of Carmen (p. 275); and reading *Don Quixote* he could split his sides with laughing (p. 59); and on receiving a letter from Martha he wrote: 'I leapt for joy—I never miss this exercise if there is the slightest reason for it' (p. 111).

And then, seriously, answering a letter from a friend who had suffered loss of a son, and writing on the anniversary of his loss of his middle daughter Sophie, he wrote: 'Although we know that after such a loss the acute state of mourning will subside, we also know we shall remain inconsolable and will never find a substitute. No matter what may fill the gap, even if it be filled completely, it nevertheless remains something else. And actually this is how it should be. It is the only way of perpetuating that love which we do not want to relinquish' (p. 386).

Lastly, what of Freud's ideas about greatness in men? 'Nor have I taken much interest in the whole species. It has always seemed to me that ruthlessness and arrogant self-confidence constitute the indispensable condition for what, when it succeeds, strikes us as greatness; and I also believe that one ought to differentiate between greatness of achievement and greatness of personality' (pp. 295-6).

One could claim that on the evidence of these letters Freud was human, and was a man of deep feeling, and he is already generally recognized as great in achievement. It may be that he could be said to have been great, too, in personality.

D. W. WINNICOTT



*The Psychology of Jung: A Critical Interpretation.* By AVIS M. DRY. London: Methuen. 1961. Pp. xiv + 329. 35s.

Dr Dry sets out to give a balanced account of Jung's work, less blindly positive than that of some of his enthusiastic supporters, and less blindly negative than that of some Freudians. In many respects she succeeds in her aim. She gives a good description of the development of Jung's thought, and of his changing relationship with Freud. She draws attention to one of his most baffling characteristics, his tendency to make extreme opposite statements at different times without seeking to reconcile them, and she is refreshingly clear about his humanist attitude to religion.

On the other hand, at times one has the impression that she has not yet had enough experience of analysis to be really capable of judging the relative merits of Freud and Jung. In a way she undervalues both. Nobody with any real insight into human nature can doubt the universality of the Oedipus complex, but she quotes uncritically various psychologists who do not accept it. At the same time, she appears not really to understand how archetypal symbolism works. People do not have 'a conscious or near-conscious perception that the desire for integration of personality can be expressed appropriately by a circle' (p. 119). The kind of thing that actually happens is that the patient painfully sacrifices some cherished attitude, without it feels almost impossible to live, realizes against his will that this has been a step forward—and then to his surprise finds himself dreaming a profoundly satisfying dream about some circular object.

One of the most interesting chapters is that in which Dr Dry relates Freud and Jung to their social backgrounds. She thinks that Jung grew up in a group that still to some extent retained the medieval assumption of a stratified society, in which everyone has both rights and duties dependent upon his rank, and that this gave him a certain basic security. Freud, on the other hand, grew up in a commercialized urban society, in which the individual was expected to succeed in competition against other individuals, and he belonged in addition to an unpopular subgroup. Thus, for example, Freud's tendency to assume that no child ever has the slightest desire to grow up is, in her opinion, a by-product of anti-semitism rather than a perception of a universal human characteristic.

J. DARROCH

*The Historical Development of British Psychiatry.* Vol. 1. 18th and 19th Centuries. By DENIS LEIGH. Oxford: Pergamon. 1961. Pp. xiv + 277. 70s.

Since the publication of Hack Tuke's *Chapters in the History of the Insane in the British Isles*, which Dr Leigh confesses he finds boring, there has been no further attempt to write a history of British psychiatry. The present book is, therefore, not only welcome but long overdue.

Dr Leigh is more interested in individuals than in concepts, so that he does not give a clear exposition of the issues which were at stake in the period under discussion. The first third of the book summarizes British psychiatric writings, but does not relate them to the development of medical thought and practice during this period. The rest of the book is devoted to biographies of Haslam, Prichard and Conolly. The chapter on Haslam firmly reinstates him as the first great European clinical psychiatrist. It is unfortunate that his association with the maladministration of Bethlem has led many medical historians to overlook his true worth. Leigh compares the reputations of Haslam and Conolly and the careers of these two men seem to have a lesson for us today. Conolly, a failure at the age of 45, luckily managed to obtain the control of a large mental hospital in time to put into effect the ideas of Gardiner Hill on 'no restraint'. From then on he became the most outstanding English psychiatrist. The young psychiatrist of today might draw the conclusion that it is better to climb on the passing bandwagon than to devote years of study to his specialty.

Although a history of British psychiatry, this book does not give an adequate account of the Scottish contribution. We are told that Clouston's textbook was clearly written under the influence of Kraepelin's work. This is rather unlikely since Kraepelin's *Compendium* and Clouston's *Clinical Lectures on Mental Diseases* were both first published in 1883. In fact Clouston always believed that he had influenced Kraepelin and that the latter's 'dementia praecox' was really the same as his 'adolescent insanity'.

There is no doubt, however, that this book will become an essential part of any psychiatric library. It is excellently produced and contains many interesting illustrations. However, is it really necessary to include illustrations of bleeding bowls, drug jars and scarifiers in a book of this kind? Perhaps their omission might have reduced the price of this book, which is rather high.

F. J. FISH

*Style in Language*. Ed. by THOMAS A. SEBEOK. London and New York: Wiley. 1960. Pp. xvii + 470. 76s.

In 1953 the Committee on Linguistics and Psychology of the Social Science Research Council sponsored a Summer Seminar which resulted in 1954 in the publication of a *Morton Prince Memorial Supplement* to the *Journal of Abnormal and Social Psychology*, outlining general research problems in what is called 'psycholinguistics'. The Committee has since sponsored several conferences on selected problems of research in this field, and the conference on literary style, from which the papers in *Style in language* are taken, is the fifth. At this conference psychologists, linguists and literary critics read papers to their own colleagues and to those in other disciplines and each group of papers was followed by an interdisciplinary discussion. For the student of linguistics there is excellent fare in abundance, but many general psychologists will want to confine their attention to the contributions of J. B. Carroll (a factor-analysis of prose style), C. E. Osgood (suicide notes), J. J. Jenkins (high and low commonality of association) and Roger Brown and Albert Gilman (single and plural second person pronouns). Also of general interest are the summings up for the psychologists (whose papers were given a cold and unenthusiastic reception) by Roger Brown and G. A. Miller. It is perhaps some consolation that the literary critics and linguists were also at cross-purposes, to a lesser extent. It is certainly regrettable that the members of six different disciplines failed to understand one another although the many problems put forward in the twenty-two papers and subsequently discussed are interesting and important, and it is to be hoped that before another conference takes place, psychologists, linguists and literary critics alike, will study some of the 462 references listed at the end of the book, especially those with which they are least acquainted.

J. W. THOMPSON

*The Biology of Art*. By DESMOND MORRIS. London: Methuen. 1962. Pp. 176. 36s.

Dr Morris is a zoologist, television personality and artist. Congo, the main subject of this book, is a chimpanzee, television personality and -artist? Dr Morris thinks yes. In his first four chapters he reviews the evidence that infra-human primates can draw and paint artistically. There are thirty-two contenders but there is adequate information about only two of these. Morris's Congo and Schiller's Alpha (whose drawings are perhaps the most interesting of the lot). Schiller and Morris have demonstrated without doubt that ape doodling is under visual control. Can more be said? Unfortunately there are wide individual differences between all the monkeys and apes, and conclusions about the origins of art must be made from what the many drawings have in common. Here the size of sample is the drawback, since the author has only the two main performers to compare. A sceptic might say that the evidence is strong regarding only two tendencies: to mark over existing figures; and to mark, often centrally, in blank spaces. The evidence that these apes also balance pictures is quite good, but the book carries examples of drawings and paintings that are not balanced. Who has the sense of balance, the ape, or he who selects the pictures? Taking the drawing situation as a whole, perhaps the author has underestimated the way in which subtle human responses may have acted as reinforcers for the experimental subjects. They may have played a part in determining the change in style which Congo showed over the months.

Dr Morris does a great service in drawing attention to Rhoda Kellogg's inaccessible and very important classification of the developmental stages in children's doodles. Do the apes show similar stages? Morris claims that Congo does, but the reader cannot judge for himself. There is of course the impossibility of including all the evidence in the book. It looks, for instance, as though the completely unsophisticated Charlie has produced (Fig. 7) a circle as good as Congo did in his maturest period, and this is said to be one of the most difficult shapes calligraphically. In the concluding chapter the author takes his case as proven and writes fluently and brilliantly on the implications. Although his approach to the origins of art is of course evolutionary.

he sees that two of the things which helped in the development of human art were its instructional and religious functions. Ape art (to which some modern art has regressed) shows what may have been happening up to this stage.

The author is right when he says that more experiments are needed; but he is to be applauded for writing such an entertaining and important account. The book is beautifully produced.

B. M. FOSS

*Clinical Process. A New Approach to the Organization and Assessment of Clinical Data.*

By E. KUNO BELLER. New York: Free Press. 1962. Pp. xx + 394. \$10.00.

The cumulatively enormous amount of endeavour that has gone into the writing of clinical notes is all too often prodigally wasted through lack of any system. For so many research purposes it is not enough to know that X suffered from *a*, *b* and *c*. It is necessary to know that he did not suffer from *d*, *e* and *f*. Without a clear system, it cannot be known whether absence of reference to *d*, *e* and *f* means that X did not suffer from them or whether he was never asked or investigated with these in view.

Dr Beller was, therefore, amply justified in making a detailed analysis of the content of clinical records at the Child Development Center, New York, to ascertain their usefulness for particular kinds of research. He found that such areas as Presenting Problems, Child-Parent and Child-Sibling Relationships were adequately covered; whereas Birth, Postnatal Development, History of the Family Group were not.

The amount and type of data provided by the different disciplines was investigated and revealed that some disciplines exceeded expectation and some fell below. An assessment was made of diagnostic and treatment procedures. Inadequate records were kept *inter alia* of what the therapist actually did in treatment; whereas initial diagnosis and subsequent attitudinal and behavioural changes were well covered. As a result of his systematic investigations the author was able to state what sorts of research could profitably be carried out with the existing records and how such records might be improved.

It would be difficult to claim that the book is fascinating reading for those not directly concerned with its particular problems; on the other hand, there are very many clinics or hospitals which should be directly concerned and which could benefit considerably from Dr Beller's careful enquiry.

G. A. FOULDS

*Introduction to Psychology.* By CLIFFORD T. MORGAN. New York and London:

McGraw-Hill. Second edition. 1961. Pp. xx + 727. 58s.

*Readings for an Introduction to Psychology.* By RICHARD A. KING. New York and

London: McGraw-Hill. 1961. Pp. xii + 388. 31s.

The sixty readings contained in the second-named book are mostly excerpts from empirical papers published by well-known psychologists in American journals of recent years. While the book is intended primarily for use along with the new edition of 'Morgan' the coverage parallels sufficiently closely that of other current introductory texts. For quality, quantity, and diversity of reading, it is good value for money, and teachers of psychology who, for any reason, want a wide-ranging pot-pourri will find it worth considering.

I. M. L. HUNTER

(Editor's note. The 'new Morgan' is little different from the old, except that the illustrations are almost entirely new. They reflect changing fashions in typography rather than any improvement in communicating information; in some cases, indeed, there seems to be a loss in clarity. An accompanying *Study Guide* has not been received.)

*Proceedings of the First International Congress on Ergonomics, Stockholm 1961.*

*Ergonomics*, Vol. 5, No. 1. London: Taylor and Francis. 1962. Pp. 336. £2. 10s.

This enlarged issue of the journal *Ergonomics* is entirely given over to contributions read at the Stockholm meeting of the International Ergonomics Association. There are in fact 51 of these papers, occasionally abstracts, by authors from 14 of the 19 countries represented by the delegates. Most of the papers are experimental.



The contributions are loosely grouped into four sections, the length of the section headings being an indication of the taxonomic difficulties in this field. The first section, 'Speed of Work and its Relation to Physiological Stress and Systems of Payment', is largely physiological. This section is composed of items on energy expenditure and working capacity, rates and techniques of work, fatigue tests, and a group of papers on heat stress. 'Systems of payment' are barely mentioned. Section 2, 'Adjustment of Work and Working Environment for Older People', is more homogeneously grouped round the problems of ageing, with the exception of a stray contribution on Finnish lumberjacks. In §3, 'Evaluation of Work and Working Environment in Ergonomic Terms', there is a collection of theoretical and experimental papers, describing both industrial and laboratory studies, on problems more 'molar' than those treated in the first section. The fourth section, 'Miscellaneous Papers', ranges from physiological to social and economic problems in a way both convincingly miscellaneous and convincingly international. The issue concludes with data on the International Ergonomics Association.

In the journal *Ergonomics*, the practice hitherto has been to request and print summaries of each paper in English, French and German. It is particularly unfortunate for English readers that this facility has been omitted in the Stockholm number, since five of the papers are in German, and three in French.

As with most conference reports, the diversity of the papers makes any general appraisal difficult, although clearly these proceedings will be consulted by persons with a special interest in any of the included topics. Considering this issue as a kind of paperbacked book, the casual reader will probably find the price discouraging.

D. H. HOLDING

*A Review of Experimental Research in Graphology, 1933-1960.* By FRITZ A. FLUCKIGER, CLARENCE A. TRIPP and GEORGE H. WEINBERG. *Perceptual and Motor Skills Monograph Supplement*, I-V 12. Pp. 24. \$1.50.

*A Table of Values for the Freeman-Tukey Square-root Transformation.* By GERALD M. MEREDITH and CONNIE C. G. WONG. Same series, I-V 13. Pp. 12. \$1.00.

*Intralist Associations Data for 99 words of the Kent-Rosanoff Word List.* By ERNST Z. ROTHKOPF and ESTHER U. COKE. *Psychological Reports Monograph Supplement*, 2-V8. Pp. 12. \$1.00.

*Measurements of Association Value (a), Rated Associations (a') and Scaled Meaningfulness (m') for the 2100 CVC Combinations of the English Alphabet.* By CLYDE E. NOBLE. Same series, 3-V8. Pp. 35. \$3.00.

The above are reprints from the journals of the Southern Universities Press (Missoula, Montana, U.S.A.), issued and sold separately as Monograph Supplements: reference to this commendable practice has already been made in this *Journal* (1961, 52, 307).

The review of work in graphology is, sensibly, limited to articles of which the authors 'tried to test one or more hypotheses, and not merely to state them and provide rationales'. Just over 100 studies, in a wide variety of languages, are considered; many positive findings are reported, but the authors call attention to the need—and the scope—for much additional fundamental research.

The Freeman-Tukey square-root transformation ' $\psi$ ' ( $\sqrt{x + \sqrt{x+1}}$ ) is a device for stabilizing variance where this requirement is not met, e.g. (to quote the authors) 'when raw data are counts of the numbers of events occurring in a fixed exposure (in space or time or both)'. Values of  $\psi$ , and of  $\psi^2$ , are tabulated for all integers between 0 and 1000.

Rothkopf and Coke present an analysis of 'intralist' associations among words contained in the Kent-Rosanoff Word List, the intention being that the data 'might be generally useful in verbal learning studies'. An example is given of the 'use' of the table, but this does not show how the information that can be extracted from the table can serve a practical purpose. No reference is made to the possibly limited validity of the 'norms' on which the analysis is based, a point touched upon by Anne Anastasi, who indeed, in her *Psychological Testing* (reviewed elsewhere in this issue) describes the Kent-Rosanoff Word List as having fallen into disuse.

Rather similar misgivings are aroused by Clyde Noble's study. The 'CVC' combinations' with which he is concerned are 'consonant-vowel-consonant trigrams', i.e. three-letter words and comparable 'paralogs' (nonsense syllables) with the vowel in the middle and the second consonant different from the first. A detailed survey of work on meaningfulness and association value of nonsense syllables is included, the author paying a handsome tribute to Ebbinghaus. Offprints received along with the monograph criticize Noble's standardization on methodological grounds. Detailed comment is outside the scope of a short review, but the inapplicability of the scaled values on this side of the Atlantic will be obvious when one notes that fourth place for meaningfulness, following MAN, WAR and CAR, is taken by TEX. Again, how long can a standardization be said to last? The author rightly calls attention to the added meaning that has accrued to such 'CVC's' as BOP, or MIG, but how can one gauge the reverse process? It would seem that maintaining a check on 'meaningfulness' could become a full-time occupation. B. SEMENOFF

*The British Journal of Social and Clinical Psychology*. Edited by MICHAEL ARGYLE and JACK TIZARD. London and New York: Cambridge University Press. Vol. 1, Part 1, February, 1962. Pp. 80. 20s. (\$3.50).

It is always a pleasure to welcome a new contemporary, the more so on the rare occasion of an addition to the British Psychological Society's family of journals. The appearance of the *Brit. J. Soc. Clin. Psychol.* marks the realization of a long-standing project: that it should have happened in Jubilee Year is gratifying and appropriate.

The present issue contains eight papers, of which four may be regarded as 'Clinical', and two as 'Social', while the remaining two (H. J. Eysenck on 'Response Set, Authoritarianism and Personality Questionnaires', and J. K. Wing on 'Institutionalism in Mental Hospitals') may be said to span both fields. All but one come from the South of England, and if in consequence there is a slight suggestion of in-breeding, this must not be taken as likely to be a constant feature. The Editorial Board includes representatives of Australia, the United States and the U.S.S.R., and (following an explicit statement of editorial policy) p. 2 of the wrapper says 'Contributions will be welcomed from psychologists of all nationalities, but must be written in English'. Apropos of the latter clause it is pleasant to be able to say that on the whole the papers are eminently readable, with an almost complete absence of the jargon that mars so much of present-day scientific writing.

Inevitably there is some unevenness of quality, but all the authors state their intentions—or specific hypotheses—very clearly. The same goes for the conclusions, even though in some cases one might question their validity. In other words, the approach is commendably empirical, and the subject-matter and treatment for the most part stimulating. Even at this early stage the 'new journal' may be regarded as having developed its own individuality. Further details of subscription rates, etc., will be found on the inside back cover of this *Journal*. B. SEMENOFF

*Productivity*. Oxford: Pergamon Press. Six issues per year. Vol. 1, No. 1. January/February 1962. Pp. 56. Annual subscription rate £3. 10s. or \$10.

It is probably unfair to judge a new journal of this type by its first issue since editorial policies—however well conceived—take some time to become effective, and potential contributors have to learn the special character which the journal wishes to adopt. Thus, presumably because of its infancy, this first issue shows little individuality to distinguish it from the already fairly numerous journals of its type intended for the desk of the business executive. The articles range from the design of business forms to new developments in lift trucks, and from management succession to the uses of hover trucks.

The style is intended for ease of reading, rather than scientific precision and detail, and while the former is always desirable, the importance of the latter in the field of professional literature for managers needs more stress than it currently receives.

It is to be hoped that this Anglo-American venture, stemming from the College of Production Technology, at Ashford, Kent, will develop a somewhat more rigorous style, while retaining the virtues of readability and width of coverage.

JOHN D. HANDYSIDE

## A TWO-FACTOR THEORY OF VIGILANCE\*

By P. D. McCORMACK

*Department of Psychology, University of Manitoba, Canada*

The findings of seven studies on reaction time are related to those of the more conventional vigilance setting, discussed with respect to existing theories of vigilance and integrated within an inhibition-motivation framework.

### I. INTRODUCTION

In vigilance experiments subjects are required to report the occurrence of a sporadically presented near-threshold signal while monitoring displays such as the Mackworth Clock Test (Mackworth, 1950) or various simulated radar and sonar devices. In such situations, number of signals detected has typically been employed as the dependent variable. Should the signal appear sufficiently above threshold to enable the subject to detect it consistently, more precise measures of performance become available. One of these is response latency or reaction time. Such a measure has been employed in a series of experiments, the results of which compare favourably with those of the more conventional vigilance setting. The reaction-time findings, however, are not consistent with various theories of vigilance. The purpose of the present paper is to point out these inconsistencies as well as to offer an alternative theoretical evaluation of the reaction-time data.

### II. THE REACTION-TIME EXPERIMENTS

In each of seven studies the subject's task was to depress a micro-switch as fast as possible each time light from a 15 W. bulb appeared for 100 msec. through an aperture 1 cm. in diameter. The distances separating the subject from the aperture ranged from 7 to 15 ft. The experiments were conducted under normal daylight illumination conditions and the experimenter was located in either a sound-deadened cubicle or a separate room. Unless otherwise stated, the intervals between stimuli were 30, 45, 60, 75 and 90 sec., all subjects experiencing each interval once every 5 min. The order in which the stimuli appeared was random.

Although four of the studies have been reported extensively elsewhere (McCormack, 1958, 1959, 1960; McCormack & Prysiazniuk, 1961), a brief description of each is necessary to enable the reader to note the resemblances between the reaction time and vigilance findings.

Experiment 1 was designed with the purpose of determining whether performance would deteriorate over time on task, a phenomenon characteristic of the signal-detection data (McCormack, 1958). Such was the case. In addition, reaction time was invariant with length of inter-stimulus interval and showed increasing amounts of improvement following rest periods of 0, 5 and 10 min. duration, respectively.

\* A portion of the research reported in this paper was supported by a grant-in-aid from the Associate Committee on Experimental Psychology of the National Research Council, Canada (Grant APBT-40). A brief account of the author's theoretical position was presented at the Second Annual Scientific Meeting of the Psychonomic Society held at Columbia University, New York, September 1961.



Experiment 2 was performed with the purpose of assessing the relative effects on reaction time of providing and withholding knowledge of results of performance (McCormack, 1959). Reaction times became progressively longer throughout the session, the slope of the function being steeper under a 'no knowledge' than under a 'knowledge of results' condition. For both treatments reaction time was a decreasing function of interstimulus interval length.

Experiment 3 provided a test of the hypothesis that the contradictory findings of the two earlier experiments with respect to interval length were due to the employment of male subjects in the former and female subjects in the latter investigation (McCormack, 1960). Ten males and ten females served as subjects in each of two sessions, which were separated by a period of one week. Since reaction time was invariant with length of interstimulus interval in each of the sessions for subjects of both sexes, the hypothesis was rejected. Reaction time increased with time on task, the slope of the function not differing reliably for male and female subjects or from session to session.

In Experiment 4 an assessment was made of the effects on reaction time of three degrees of variability of interstimulus interval (McCormack & Prysiaziuk, 1961). The light appeared on the average once per min. but the degree of regularity with which it was presented differed from day to day. In one condition it appeared every 60 sec., in another the standard intervals were employed, while in the remaining condition the stimuli were separated by 10, 35, 60, 85 and 110 sec. intervals. As in previous experiments reaction time became progressively longer with time on task, the slopes of the functions for the three treatments not differing reliably from one another. However, the more regular the appearance of the stimulus the faster were the overall mean reaction times. An invariant relation was again observed between reaction time and length of interstimulus interval.

Experiment 5 was undertaken with the purpose of re-examining the effects on reaction time of providing and withholding knowledge of results of performance (McCormack, Binding & Chylinski, 1961). The first part of the study was a replication of Expt. 2 with twelve male and eight female subjects participating in a 35 min. session on each of two consecutive days. Under the 'knowledge' condition a red or a green pilot light was illuminated immediately following each response. The red light indicated a reaction time which was slower than the previous one while the green light signalled a faster response. These coloured lights were not employed in the 'no knowledge' treatment. In the second part of the experiment, fourteen male and six female subjects served in a single 35 min. session under a 'no knowledge' condition in which the two pilot lights were presented simultaneously following each response. For both of the 'no knowledge' treatments reaction times increased with time on task and decreased with length of interstimulus interval. Under the 'knowledge' condition, however, performance remained invariant with both task duration and interval length.

The same conditions were employed in Expt. 6, but in this study all subjects were given each of the three treatments (McCormack *et al.* 1961). The performance of thirteen male and eleven female subjects did not differ reliably from that of the forty subjects of Expt. 5. In the light of the findings of Expts. 5 and 6 as well as of the highly questionable status of those of Expt. 2, the evidence for which has been discussed elsewhere (McCormack & Prysiaziuk, 1961), it must be concluded that the provision of knowledge of results of performance prevents any systematic increase in reaction time during the course of an experimental session.

Experiment 7 was designed to study the effects on performance, if any, of the sudden introduction of an extraneous stimulus (Gordon, 1961). Nine male and seven female subjects served for 40 min. on each of two consecutive days. One of the treatments involved the introduction of an intense auditory signal following the 30th min. of the task. Performance improved following the presentation of this signal although the change could not be demonstrated to be statistically dependable. During both sessions, reaction times increased over time and were invariant with length of interstimulus interval.

An analysis of the 'no knowledge' data of Expts. 1-7 reveals a linear relation between reaction time and task duration. The slope of this function, which holds up reasonably well from experiment to experiment, is  $4.43t$  where  $t$  represents successive 5 min. time blocks. Research is currently being planned with the purpose of determining those conditions which do and those which do not produce dependable changes in the slope parameter.

## III. THE REACTION-TIME FINDINGS AND THEORIES OF VIGILANCE

The reaction-time and signal-detection findings—the latter have been extensively summarized elsewhere (McGrath, Harabedian & Buckner, 1959)—are comparable in several respects. In both cases performance deteriorates progressively throughout a session under 'no knowledge' conditions, shows no systematic change over time when knowledge of results is provided, improves following periods of interpolated rest and is either an increasing function of or invariant with length of interstimulus interval. Only where the signal has been presented in a regular fashion do the data of the two situations conflict, reaction time becoming progressively longer, whereas probability of signal detection is invariant with time on task.

Because of the forementioned similarities, the reaction-time findings should be reasonably consistent with any theory of vigilance which has been devised to account for the signal-detection data. A number of such theoretical models have been proposed (Frankmann & Adams, 1960).

Mackworth (1950) has postulated that in the absence of reinforcement of the response inhibition will accumulate, resulting in a systematic deterioration of performance. Provision of knowledge of results serves as the reinforcer, while the introduction of an extraneous stimulus will allow the inhibition to dissipate. Consistent with the Mackworth model is the observation that, whenever knowledge of results of performance is withheld, reaction times show a progressive lengthening with time on task (Expts. 1-7). Also accounted for are the findings that reaction time improves following the presentation of an intense auditory signal (Expt. 7) and is constant over time when knowledge of results of performance is provided (Expts. 5 and 6).

Holland (1958) and Bakan (1959), like Mackworth, employ the concept of reinforcement to account for certain of the signal-detection findings. For both, the reinforcing event is the detection of a signal which maintains responses of orienting toward and fixating or scanning the display. Holland calls these 'orienting responses' while Bakan prefers the term 'attentive behaviour'. Since in a reaction-time setting the stimuli are detected 100% of the time, these theoretical models are of limited applicability.

According to Broadbent (1953, 1958), whenever the task stimuli lose their initial novelty, i.e., become more and more similar to the immediately preceding ones, attention will fluctuate and performance will be impeded. When the subject is confronted by stimuli which are irrelevant to the task, as in the case when an extraneous signal is presented, or when knowledge or rest periods are introduced, the task stimuli regain their initial novelty and performance improves. Berlyne (1951) has incorporated the novelty concept into a general theory of attention and perception which, in its present form, is not directly applicable to either the signal detection or reaction-time findings. Similar to the Broadbent model is one devised by Scott (1957), where lack of stimulus variation rather than reduced novelty is postulated as the determiner of performance decrement. The following findings are consistent with both the Broadbent and Scott models: reaction time increases over time under 'no knowledge' treatments (Expts. 1-7), is constant throughout a session when knowledge of results of performance is provided (Expts. 5 and 6), and improves following periods

of interpolated rest, as well as following the presentation of an extraneous stimulus (Expts. 1 and 7). At variance with both theories, however, is the finding that reaction time will become longer over time under a 'no knowledge' condition where additional stimulation is provided following the elicitation of a response (Expts. 5 and 6).

The most recently elaborated theory of vigilance, and one that has received considerable experimental attention, is an extension by Baker (1959) of a set of hypotheses put forward earlier by Deese (1955). According to Baker, the subject's expectancy is high, hence his performance optimal, when he is allowed an accurate perception of the sequence of stimuli. Such is presumably the case when knowledge of results of performance is provided or when the stimulus appears regularly. Baker also postulates that, once the stimulus sequence is known, expectancy will increase with length of interstimulus interval. The Baker model is the only one of those previously discussed which is able to account for faster reaction times under conditions of regular stimulus presentation (Expt. 4) as well as following the longer interstimulus intervals (Expts. 2, 5 and 6). Also consistent with the theory are the observations that reaction time increases throughout a 'no knowledge' session (Expts. 1-7) and is constant over time when knowledge of results of performance is provided (Expts. 5 and 6). At odds with the expectancy model, however, are the findings that reaction time is invariant with interstimulus interval length (Expts. 1, 3-7), and is an increasing function of time on task under 'no knowledge' conditions where the stimuli appear regularly (Expt. 4), or when pilot lights are illuminated following each response (Expts. 5 and 6).

#### IV. AN INHIBITION-MOTIVATION THEORY OF VIGILANCE

An examination of the reaction-time and signal-detection findings, in conjunction with various theories devised to account for the latter, reveals numerous gaps between theory and experimental results. For most of the models there are data which are consistent, neutral and at variance with respect to their implications. In addition, in many cases their structure is such that predictions generated from them are of an equivocal nature. For example, the author finds it extremely difficult to see how the Baker model is consistent with the most general finding of all, that performance deteriorates with time on task under conditions in which knowledge of results of performance is withheld. For these reasons, an attempt will be made to integrate in a rigorous fashion all of the data obtained in the reaction-time setting only, in the hope that as more information is accumulated the theory may be successfully extended beyond its initial boundary conditions to include vigilance phenomena of a more general nature. To facilitate a comparison of theory and data, a summary of the reaction-time findings is presented in Table 1.

The observations that reaction time is a linearly increasing function of time on task under 'no knowledge' conditions, but is invariant with task duration when knowledge of results of performance is provided, create a strong argument for a general reinforcement interpretation of the data. If it may be assumed, as Mackworth (1950) already has, that knowledge of results reinforces whatever responses are necessary for the maintenance of optimal performance, then the temporal course of reaction



time under both 'knowledge' and 'no knowledge' conditions may be accounted for by postulating the development of inhibition in the absence of reinforcement of the response. This inhibition construct (*In*) is strikingly similar to that employed by Spence (1960) to account for the findings of certain instrumental conditioning data; for example, phenomena characteristic of partial reinforcement and extinction procedures.

Table 1. *A summary of the reaction-time (RT) findings*

Expt.	The time variable	The interval variable	Miscellaneous findings
1	Increase in RT	No change in RT	Decrease in RT following rest periods
2	Increase in RT	Decrease in RT	Increase in RT over time less pronounced in a 'knowledge' condition
3	Increase in RT	No change in RT	No differences between sexes
4	Increase in RT	No change in RT	Faster mean RT under conditions of more regular stimulus presentation
5a and 6a (no knowledge and stimulation control)	Increase in RT	Decrease in RT	—
5b and 6b (knowledge)	No change in RT	No change in RT	—
7	Increase in RT	No change in RT	Decrease in RT following an auditory signal

The improvement in reaction time, following the sudden and unexpected presentation of an extraneous stimulus as well as following periods of interpolated rest, is consistent with the notion that *In* dissipates under such conditions, an assumption not unlike that employed by Pavlov (1927) to account for the recovery of conditioned responses following the introduction of extraneous stimuli and rest periods. If it is assumed additionally that, under certain as yet unspecifiable conditions, some fraction of *In* will dissipate between the presentation of stimuli, this accounts for the decreasing function relating reaction time and length of interstimulus interval which has been occasionally observed under 'no knowledge' treatments. Also consistent with this postulate are the findings of the 'knowledge' portions of Expts. 5 and 6, where reaction time was invariant with both task duration and interval length.

The relations existing between the inhibition construct and each of the forementioned experimental variables may be summarized as follows:

$$In = f(t, 1/r, 1/e, 1/i, x_1, \dots, x_n),$$

where *t* = time on task, *r* = amount of rest, *e* = intensity of extraneous stimulus, *i* = length of interstimulus interval, *x*<sub>1</sub>, ..., *x*<sub>*n*</sub> = unknown variables. Note that the intervening variable *In* is defined, as the remaining construct will be, in purely quantitative terms as a mathematical function of certain experimental variables and that no mention is made of its possible locus, structure or functioning in the nervous system. The author's preference for this approach is based on the argument that any surplus meaning that might be associated with these added physiological connotations does not improve either the integrative or predictive value of the model (Spence, 1956).

Since the inhibition postulates are unable to account for the faster overall mean reaction times observed under conditions in which the stimuli are presented in a relatively regular fashion, a case may be made for a second theoretical construct.

Since it is conceivable that variations along a regularity dimension will affect the drive level of the subject we will call this second intervening variable  $D$  in keeping with terminology employed by investigators of motivational phenomena and will define it as follows:  $D = f(1/v, y_1, \dots, y_n)$ , where  $v$  = degree of variability of the stimulus,  $y_1, \dots, y_n$  = numerous other variables, as yet unknown, which affect performance by either raising or lowering the subject's drive level.

We have now related a number of experimentally manipulable variables to each of two theoretical constructs. The only remaining task is to show how these constructs combine with each other to determine performance. Before doing so, it would be worth while to include a third intervening variable in order to account for the large differences observed between individuals at the outset of a session. We will call this third construct  $C$ , to reflect the influence on reaction time of both innate and previously acquired characteristics of the organism, and will assume that it operates as a constant throughout the duration of an experimental session. We may now state the way in which our three intervening variables combine in determining performance:

$$\text{Reaction time} = f(CD - Ia).$$

The above set of postulates successfully accounts for all of the reaction time as well as the majority of the signal detection findings reported up to the present time. These postulates may be summarized briefly in the form of sets of equations which relate a number of independent variables to theoretical constructs which in turn are related to each other and then to performance. The theory not only integrates the data accumulated so far but, of equal importance, is capable of generating predictions which are reasonably exact and free from ambiguity. Above all, the model is flexible and may be modified from time to time to account for findings which are either at variance or neutral with respect to its implications.

#### REFERENCES

- BAKAN, P. (1959). Extraversion-introversion and improvement in an auditory vigilance task. *Brit. J. Psychol.* **50**, 325-32.
- BAKER, C. H. (1959). Towards a theory of vigilance. *Canad. J. Psychol.* **13**, 35-42.
- BERLYNE, D. E. (1951). Attention, perception and behavior theory. *Psychol. Rev.* **58**, 137-46.
- BROADBENT, D. E. (1953). Classical conditioning and human watch-keeping. *Psychol. Rev.* **60**, 331-9.
- BROADBENT, D. E. (1958). *Perception and communication*. London & New York: Pergamon.
- DEESE, J. (1955). Some problems in the theory of vigilance. *Psychol. Rev.* **62**, 359-68.
- FRANKMANN, J. P. & ADAMS, J. A. (1960). Theories of vigilance. *USAF Res. and Develop. Command, Tech. Note AFCCDD-TN-60-25*.
- GORDON, A. E. (1961). Personal communication.
- HOLLAND, J. G. (1958). Human vigilance. *Science*, **128**, 61-67.
- MACKWORTH, N. H. (1950). Researches on the measurement of human performance. *Med. Res. Council, Special Report*, No. 268. London: H.M. Stationery Office.
- MCCORMACK, P. D. (1958). Performance in a vigilance task as a function of inter-stimulus interval and interpolated rest. *Canad. J. Psychol.* **12**, 242-6.
- MCCORMACK, P. D. (1959). Performance in a vigilance task with and without knowledge of results. *Canad. J. Psychol.* **13**, 68-71.
- MCCORMACK, P. D. (1960). Performance in a vigilance task as a function of length of inter-stimulus interval. *Canad. J. Psychol.* **14**, 265-8.
- MCCORMACK, P. D., BINDING, F. R. S. & CHYLINSKI, JOAN (1961). The effects on reaction-time of knowledge of results of performance. (Unpublished.)

- McCORMACK, P. D. & PRYSIAZNIUK, A. W. (1961). Reaction-time and regularity of inter-stimulus interval. *Percept. Mot. Skills*, **13**, 15-18.
- MCCRATH, J. J., HARABEDIAN, A. & BUCKNER, D. N. (1959). Human factor problems in anti-submarine warfare. *Tech. Report*, no. 1. Review and critique of the literature on vigilance performance. Los Angeles: Human Factors Research, Inc.
- PAVLOV, I. P. (1927). *Conditioned reflexes* (trans. by G. V. Anrep). New York: Oxford University Press.
- SCOTT, T. H. (1957). Literature review of the intellectual effects of perceptual isolation. *Defence Research Board of Canada. Report*, no. HR 66.
- SPENCE, K. W. (1956). *Behavior theory and conditioning*. New Haven: Yale University Press.
- SPENCE, K. W. (1960). *Behavior theory and learning*. Englewood Cliffs: Prentice-Hall.

*(Manuscript received 26 July 1961)*





## GSR CONDITIONING AND PSEUDOCONDITIONING

By IRENE MARTIN

*Institute of Psychiatry, University of London*

Subjects were placed into conditioning or pseudoconditioning groups, and received one, three, or five UCS trials (paired with CS or single). The sequence of the schedule was: ten light stimuli—tones (the UCS)—ten light stimuli. Significantly more GSR's were given to the final light series than to the first. Conditioning groups together gave significantly more responses than pseudoconditioning groups during the final lights. However, the effect of different numbers of UCS trials was more clear-cut between pseudoconditioning groups.

Basal skin resistance was significantly lower in all groups following the tones. There was no difference between conditioning and pseudoconditioning groups in this measure. Number of GSR's and basal skin resistance were not significantly correlated in the first light series, but were highly (negatively) so in the final light series. Autonomic Lability Scores to tone 1 correlated significantly with number of GSR's to both light series only in the conditioning groups. This effect was interpreted as a generalized responsiveness of these groups plus a small facilitatory effect from the extent of the S's response to the tones.

The results are discussed from two points of view: (i) physiological, and considering how various C.N.S. levels might affect GSR amplitude and threshold via excitability of spinal internuncial neurones, and (ii) within behavioural theories of learning, and considering the effect of the UCS upon level of 'drive'.

It has frequently been observed in GSR experiments that the introduction of a strong stimulus (UCS), either paired with a weak stimulus (CS) as in classical conditioning or occurring alone as in pseudoconditioning, is very rapidly followed by an increased responsiveness to the CS. The present study undertook to quantify this observation by comparing conditioning and pseudoconditioning procedures in matched groups.

The position of the pseudoconditioning factor within conditioning theory is still unclear. Conditioning experiments often include two or three initial UCS trials to eliminate 'sensitized' subjects, on the assumption that 'pure' association is the basic factor in conditioning theory and that pseudoconditioning (or sensitization as it is sometimes called) is an irrelevant, non-associative factor. From the practical point of view alone it is doubtful if such checks are effective or worth while. Commonly only two or three UCS trials are given for checking purposes, whereas many more may be needed to produce the maximal sensitization effect with certain types of responses. It has been well demonstrated with eyeblink responses and finger withdrawal to shock that a long series of unpaired UCS is capable of facilitating responses to the CS (Grant, 1943; Harris, 1941) and this factor is presumably also operating when the UCS is paired with the CS as in normal conditioning procedures. It is simply the case that in conditioning experiments the two effects—association and sensitization—are confounded.

This is not to imply that from the theoretical point of view they are necessarily conceptually different. Wickens & Wickens (1942) have argued that the phenomena of pseudoconditioning often closely parallel those of conditioning, and by taking the stimulus generalization hypothesis to a bold extreme they are able to adopt the unassailable position that associative elements are present in pseudoconditioning experiments, if only by virtue of the similarity of the external environment in which the

CS and UCS occur. If this view is correct, the pseudoconditioning factor should be encompassed within the quantitative laws governing generalization gradients. At the moment such quantitative data are lacking.

This kind of interpretation, with its emphasis on associative factors, follows the traditional Hullian view that habit strength develops with increasing numbers of paired trials—and there is an implicit belief that drive is essentially constant in conditioning situations. But other authors (Kimble, Mann & Dufort, 1955) have postulated that it is increasing drive strength which is reflected in the rising acquisition curve. Their evidence (although unconfirmed by Goodrich, Ross & Wagner, 1957) is that there was very little difference in acquisition curves of conditioned eye-blink responses between a group regularly given paired trials, and two other groups in which the UCS was presented singly on a number of trials. Yet a third view of pseudoconditioning, proposed by more physiologically oriented workers, is that it arises through sensitization in the true physiological sense, i.e. through a reduction in neural response thresholds (Davis, 1930; Hagbarth & Kugelberg, 1958). This view, too, has its opponents who offer contrary evidence (Grant, 1943). With this much theoretical uncertainty, and with so few confirmed findings, it would seem particularly necessary to concentrate at the present time on the *quantification* of the two effects under discussion in appropriately selected experimental conditions.

Of all the ideas on the nature of pseudoconditioning which have been presented, that of Davis (1930) seems particularly apposite to our present purpose since it was based on GSR experiments. Davis found that GSR facilitation was related to the effect of the UCS upon basal levels of skin resistance. It has since been further demonstrated that basal levels of skin resistance are related to response probability under certain conditions (Davis, Buchwald & Frankmann, 1955; Martin, 1960*a, b*), and an interpretation offered in terms of alterations of neural thresholds, possibly peripherally located.

The major aims of this experiment, then, were twofold. The first was to compare the facilitative effect of conditioning and pseudoconditioning procedures, using different numbers of trials. For this purpose subjects were placed into one of six groups, three receiving conditioning and three pseudoconditioning procedures, and each matched pair receiving 1, 3, or 5 UCS trials (paired with CS for conditioning and occurring singly for pseudoconditioning). The effects of the different treatments were assessed by comparing GSR's given to two series of light stimuli, one before and one after the UCS tones. The second major aim was to examine the relationship of (i) level of skin resistance, and (ii) GSR amplitude to tones, with the facilitation processes.

#### METHOD

*Subjects.* The subjects were 112 male adolescent industrial apprentices from a large engineering works; they were unacquainted with psychological experimental procedures.

*Apparatus.* The galvanometer consists of a valve voltmeter, the deflexions of which are directly proportional to changes in apparent resistance. In this instrument, four valves are employed in two pairs, in the form of an electronic see-saw which is normally held in a state of equilibrium but which is disturbed by changes in the subject's resistance. The 'active' valves comprise a high input resistance long-tailed pair, R-C coupled to a pair of output triodes. The recorder is connected between cathodes to give a low output impedance. Calibration resistors are provided to give mid-scale deflexions on each of twelve ranges, giving an overall coverage up to 1 M $\Omega$ . An almost constant current of 10  $\mu$ A. is passed across the electrodes.



The electrode system proposed by Malmö & Davis (1961) was adopted; this employs three silver-silver chloride electrodes, of which two are reference electrodes and attached to the volar surface of the forearm (resistance checks are made to ensure a reading of below 5K between them) and the third is the active electrode (in this case  $1\frac{1}{2} \times 1$  cm.) and attached to the left palm. These silver-silver chloride electrodes were used in conjunction with a 1% NaCl electrolytic solution.

The subjects were seated in a dimly lit, sound-proof room. The experimenter and apparatus were in an adjoining room. An intercommunication system between the rooms made it possible to monitor the records for movements, sighs, and other artefacts.

*Procedure.* The S's were assigned to the six experimental groups on the basis of two factors: (i) the number of GSR's given to the first series of light stimuli, and (ii) basal skin resistance during this period. Thus the six groups were matched for these two measures.

The subject was seated comfortably before a screen ( $20 \times 18$  in.) in which was fixed a small red light (the CS) provided by a 6 V., 0.3 amp. bulb. The UCS—a tone of 980 cyc./sec., 110 db. approximately half-second in duration—was delivered through headphones. The electrodes were attached to the subject and assurances given that no shocks would occur. He was requested to keep his eyes open all the time, and simply told that light and tone stimuli would be given to him and that all he had to do was to remain as still and relaxed as possible.

After a few minutes rest, a series of ten light stimuli, each 5 sec. in duration, was presented at irregular intervals ranging from 15 to 45 sec. The tones (1, 3, or 5 in number) were then given at 45–60 sec. intervals, singly in the case of pseudoconditioning groups, and paired with the light stimuli in conditioning groups (both stimuli terminating simultaneously). After an interval of approximately 40–50 sec. following the last tone, the final series of ten light stimuli was given to the subject, also at irregular intervals. Thus the arrangement of the groups was as follows:

	<i>N</i>	<i>Conditioning</i>		
Group 1 C	19	10 light stimuli	$\left\{ \begin{array}{l} 1 \text{ paired trial} \\ (\text{light} + \text{tone}) \\ 3 \text{ paired trials} \\ 5 \text{ paired trials} \end{array} \right\}$	10 light stimuli
Group 3 C	19			
Group 5 C	18			
<i>Pseudoconditioning</i>				
Group 1 PsC	19	10 light stimuli	$\left\{ \begin{array}{l} 1 \\ 3 \\ 5 \end{array} \right\}$	tone (unpaired)    10 light stimuli
Group 3 PsC	19			
Group 5 PsC	18			

## RESULTS

Five measures were obtained from the skin-resistance records of all subjects. The first two were counts of responses occurring (i) to the first series of lights, and (ii) to the final series of lights. To qualify as a response, a deflexion of greater than 1K must occur within 5 sec. of the light onset. This criterion was selected since the CS-UCS interval was 5 sec. in the conditioning groups, and any lengthier criterion might have included incidental 'spontaneous' responses. The next two measures were of mean basal skin resistance during the first and final light series. These readings were obtained from the ten pre-stimulus (light) readings of skin resistance, and transformed into log scores. The fifth measure was an autonomic lability score based on reactions to the tones (Lacey, 1956).

Table 1 shows the results of the analysis of variance on responses to the pre- and post-tones light series (measures (i) and (ii)). The between-groups variance is, of course, non-significant, since the groups were equated on the number of responses given to the first light series. The increase in responses from the first to the final light

series is highly significant, and presumably reflects some process of facilitation. The groups  $\times$  trials variance is also significant, and a further analysis of this result by means of 't' tests gives the results shown in Table 2. There is clearly a greater overall effectiveness from paired light-tone trials than from unpaired tone trials. This breakdown of results also shows that three and five pseudoconditioning tones are more effective than one in facilitating responses to light; and, to a less significant extent, five pairings of light-tone are more effective than one.

Table 1. *Analysis of variance on GSR's to light series (pre- and post-tones)*

Main effects	D.F.	S.S.	M.S.V.	F	P
Between groups	5	58.49	11.698	0.746	N.S.
Between light series (pre- and post-tones)	1	110.04	110.04	37.026	< 0.01
Interactions					
Groups $\times$ series	5	67.42	18.484	4.537	< 0.01
Between people within groups	106	1662.39	15.683	5.277	< 0.01
Residual	106	315.05	2.972	—	—
Total	223	2213.39			

Table 2. *'t' tests carried out on A/V in Table 1*

(All 't's between groups\*, final light series.)

Between: 3 conditioning groups and 3 pseudoconditioning groups (final light series)	't' = 4.6532	P < 0.01
Between: Group 5 PC and group 3 PC	't' = 0.6858	N.S.
Group 5 PC and group 1 PC	't' = 2.8204	< 0.01
Group 3 PC and group 1 PC	't' = 2.1649	< 0.05
Between: Group 5 C and group 3 C	't' = 0.8198	N.S.
Group 5 C and group 1 C	't' = 2.3048	< 0.05
Group 3 C and group 1 C	't' = 1.506	N.S.

\* Groups 5 PC, 3 PC, and 1 PC received 5, 3, and 1 pseudoconditioning trial, respectively. Groups 5 C, 3 C, and 1 C received 5, 3, and 1 conditioning trial, respectively.

Table 3 gives the results from basal skin resistance measurements. There is a significantly *lower* level of skin resistance following both paired and unpaired tones, and in this case no differential effect is observed between conditioning and pseudoconditioning groups, i.e. the groups  $\times$  light series interaction is not significant.

The next section of the data deals with autonomic lability scores. These were obtained on the basis that the tones might have more or less impact on the subjects and so affect the degree of subsequent conditioning or pseudoconditioning. While all the correlations are in the expected direction—those Ss who respond with bigger amplitude to the tones give more responses to the final light series—only two of the correlations reach an acceptable level of statistical significance (see Table 4). These indicate that in the three conditioning groups there is a positive relationship between reactivity to tone 1 and number of responses to final light stimuli. It cannot be argued, however, that facilitation to the lights was produced by excessive reactivity to the tone, for the second significant correlation was, unexpectedly, between the autonomic lability score to tone 1, and the number of responses given to the *first* series of lights, i.e. prior to auditory stimulation.

Two other correlations of interest were calculated: one between basal skin resistance and number of responses during the first light series, where  $r = -0.1408$  (N.S.) and the second between basal skin resistance and number of responses during the final light series, where  $r = -0.5249$  (significant at the 1 % level). Apparently, under these experimental conditions, it is only *after* the UCS tones that level of basal skin resistance significantly affects response probability to the lights.

Table 3. *Analysis of variance on basal skin resistance during light series (pre- and post-tones)*

	D.F.	S.S.	M.S.V.	F	P
Main effects					
Between groups	5	0.906345	0.181269	0.989	N.S.
Between light series (pre- and post-tones)	1	0.141002	0.141002	11.728	< 0.01
Interactions					
Groups $\times$ light series	5	0.041810	0.008362	0.696	N.S.
Between people within groups	106	19.425637	0.183261	15.243	< 0.01
Residual	106	1.274403	0.012023	—	—
Total	223	21.789197			

Table 4. *Correlations between autonomic lability scores and number of GSR's given to light stimuli*

A.L.S. with number of GSR's: 1st light series

	$r$	P
All pseudoconditioning groups (tone 1)	-0.2186	N.S.
All conditioning groups (tone 1)	-0.3121	< 0.05
Groups 3 PC and 5 PC (tone 3)	-0.0649	N.S.
Groups 3 C and 5 C (tone 3)	-0.1780	N.S.

A.L.S. with number of GSR's to final light series

All pseudoconditioning groups (tone 1)	-0.2084	N.S.
All conditioning groups (tone 1)	-0.4937	< 0.01
Groups 3 PC and 5 PC (tone 3)	-0.0616	N.S.
Groups 3 C and 5 C (tone 3)	-0.2112	N.S.

### DISCUSSION

The findings are clear in showing that the probability of occurrence of a GSR to a weak stimulus is increased after several intense stimuli, and that this effect is associated with a reduction in basal skin resistance. These results supplement Davis's (1930) data on facilitation, which were expressed in terms of increased amplitude of response and also associated with a lowered skin resistance to the tones. At that time he offered a tentative explanation of the inter-relationships based on the facts of electrotonus (the polarization of cell membranes when a direct current is passed through them), and tended to favour the notion that the effector organ might play an important part in facilitatory effects. In a later study (Davis *et al.* 1955) the authors suggest that in a state of, for example, positive electrotonus, apparent tissue resistance is increased, the stimulus threshold is raised, and the electrical responses when they do occur are larger.

But although GSR characteristics sometimes depend closely upon the state of the skin tissues prior to the application of the stimulus, there are a number of experimental conditions where these relationships are not clear. In the present experiment



the number of GSR's is clearly and significantly related to basal skin resistance only during the final light series and not during the first. Physiological evidence favours the view that threshold changes can occur at many sites within the C.N.S., and it seems unlikely that change in peripheral membrane permeability is the only factor responsible for alterations in response threshold.

There is unfortunately no direct experimental evidence on the state of the sweat glands, or of other parts of the autonomic nervous system during these experimental conditions, on which physiological argument could be based. The most relevant data are probably those provided by Wang (1957, 1958), whose experimental work and reviews indicate that the GSR is mediated by a number of facilitatory and inhibitory influences operating at different levels within the C.N.S. But it ought perhaps to be emphasized that his experimental conditions and methods (electrical stimulation and lesions) were very different from those commonly employed in psychological experiments, and they suffer from the basic disadvantage that they were not made under normally occurring physiological conditions.

Hagbarth & Kugelberg (1958) maintain that it is only in intact organisms, where the spinal reflex centre is subjected to the influence of the normally functioning brain, that the *plasticity* of reflex functions becomes evident and can best be studied. These plastic properties cannot be accounted for by spinal function alone: they depend instead upon cerebral outflows which can adjust the excitability of spinal reflex centres. 'Evidently, a sensory block, evoked by habituation, may be cancelled by the sensitization process, which strives to increase the excitability within the spinal reflex centre.' Wang essentially agrees with this position. In the case of the GSR he believes that excitatory and inhibitory impulses from the periphery and from the suprasegmental nervous structures act on the preganglionic sympathetic sudomotor neurones through the intermediary of the internuncial pool in the spinal cord, and that the algebraic sum of such influences is directly reflected in GSR amplitude. It seems a short step to the inference that GSR sensitization may similarly be determined at spinal levels.

Aside from physiological speculation, the present findings can also be considered within current psychological learning theories. They are compatible with the view that conditioning and pseudoconditioning effects simply differ quantitatively. But it is surprising (in view of what we know of gradient slopes and the importance of CS-UCS temporal contiguity) that the pseudoconditioning and conditioning effects are so similar if an explanation is considered in terms of stimulus generalization.

An hypothesis based on changes in drive strength would seem more in line with current thought, e.g. Spence (1956) on the 'hypothetical emotional response' produced by the UCS. It could well be that such a response occurs initially as widespread autonomic activation and sensitization, perhaps involving a lowering of response thresholds, and that it subsequently habituates. The function relating number of UCS trials to autonomic activation and facilitation is not yet known: the present pseudoconditioning data, which show a clear effect due to different numbers of UCS trials, represent a beginning in this direction. It would be predicted that this increase eventually reaches a maximum following which a decremental (adaptation) effect occurs. The inference from this is that 'drive' (or some allied activating factor) does not remain constant throughout a stimulus series.

Superimposed upon these changing facilitatory and adaptive effects is the additional factor of habit produced by pairing the CS with UCS. However, the effect of 1, 3, or 5 paired trials upon extinction is not a very distinct one—a finding which offers little support to the view that habit strength is a direct function of the number of paired trials.

Finally, the data obtained from the autonomic lability scores show that reactivity to tone 1 and responsiveness to both series of lights is significantly correlated in the conditioning groups. This would seem to imply a general factor of responsiveness to tones and lights, supplemented, perhaps, by a small facilitatory effect produced by the extent of the S's response to the tones, since the correlation increases from  $-0.312$  (first series) to  $-0.493$  (last series)—see Table 4.

The writer is indebted to the Bethlem Royal Hospital and the Maudsley Hospital Research Committee for a grant which made this study possible.

#### REFERENCES

- DAVIS, R. C. (1930). Factors affecting the galvanic reflex. *Arch. Psychol.*, N.Y., no. 115.
- DAVIS, R. C., BUCHWALD, A. M. & FRANKMANN, R. W. (1955). Autonomic and muscular responses, and their relation to simple stimuli. *Psychol. Monogr.* no. 405, 69, no. 20.
- GOODRICH, K. P., ROSS, L. E. & WAGNER, A. R. (1957). Performance in eyelid conditioning following interpolated presentations of the UCS. *J. Exp. Psychol.* 53, 214–217.
- GRANT, D. A. (1943). The pseudo-conditioned eyelid response. *J. Exp. Psychol.* 32, 139–49.
- HAGBARTH, K. E. & KUGELBERG, E. (1958). Plasticity of the human abdominal skin reflex. *Brain*, 81, 305–18.
- HARRIS, J. D. (1941). Forward conditioning, backward conditioning, pseudo-conditioning and adaptation to the conditioned stimulus. *J. Exp. Psychol.* 28, 491–502.
- KIMBLE, G. A., MANN, L. I. & DUFORT, R. H. (1955). Classical and instrumental eyelid conditioning. *J. Exp. Psychol.* 49, 407–17.
- LACEY, J. I. (1956). The evaluation of autonomic responses: toward a general solution. *Ann. N.Y. Acad. Sci.* 67, 123–64.
- MALMO, R. B. & DAVIS, J. F. (1961). A monopolar method of measuring palmar conductance. *Amer. J. Psychol.* 74, 106–13.
- MARTIN, IRENE (1960a). Variations in skin resistance and their relationship to GSR conditioning. *J. Ment. Sci.* 106, 281–7.
- MARTIN, IRENE (1960b). The effects of depressant drugs on palmar skin resistance and adaptation. Ch. 8 in: *Experiments in Personality*, Vol. 1, pp. 197–220. Ed. H. J. Eysenck. London: Routledge and Kegan Paul.
- SPENCE, K. W. (1956). *Behaviour theory and conditioning*. New Haven: Yale University Press.
- WANG, GING HSI (1957). The galvanic skin reflex. A review of old and recent works from a physiologic point of view. Part 1. *Amer. J. Phys. Med.* 36, 295–320.
- WANG, GING HSI (1958). The galvanic skin reflex. A review of old and recent works from a physiologic point of view. Part 2. *Amer. J. Phys. Med.* 37, 35–57.
- WICKENS, D. D. & WICKENS, CAROL D. (1942). Some factors related to pseudo-conditioning. *J. Exp. Psychol.* 31, 518–26.

(Manuscript received 7 September 1961)





## RELEVANCE AND CATEGORY SCALES OF JUDGEMENT

By ROBERT S. DAVIDON

*Bryn Mawr College*

Once adaptation levels had been established in category judgements of the height and lightness of a series of grey blocks, a tall, light block was introduced, and presented with each of the series. Half of the subjects were instructed to reject any block which was lighter than it, and half to reject only those taller. After each decision both groups categorized height and lightness as before. Shifts in scales of judgement were associated with differences in the *relevant* attributes of the standard.

### I. INTRODUCTION

Very tall or short, dark or very light-skinned, a man before us can be categorized reliably and with ease—yet categories may change. Noting that he is exceptional in height or pigmentation may affect the range to which one applies an extreme descriptive category, and may influence whom one henceforth classifies as medium. It is difficult to conceive, however, of this meeting with one person—with his thousands of different physical and behavioural characteristics—influencing untold numbers of subjective scales of judgement within the single observer. Perhaps the effects of each instance upon the observer's categories of relative judgement are not absolute for each variable, but are limited by a selective process.

Can it be demonstrated that some but not all of an object's attributes have effects on category scales of judgement? If so, what determines whether or not an effect occurs, which attributes of the perceived object have lasting consequences, and the degree to which each scale is affected?

In the laboratory the study of scales of judgements has been quantified and in some ways made more simple. Brass weights, pure tones, or paper rectangles have been judged more often than men, and usually the items of the series have been varied in but a single dimension. The typical value of each category of the scale of judgement has been accurately predicted from the magnitude of the items presented, one by one, and from the value of the background (Helson, 1947; Michels & Helson, 1949; Michels, 1953; Michels & Doser, 1955). In these studies responses were almost entirely a function of stimuli and background; there was little 'residual' variability, attributable to past experience and other inner determinants. Some effects of previous experiences (Nash, 1950) and personal factors (Ball, 1953) have been studied, but the 'residual' has been conceived in terms of states of the organism, essentially cumulative, rather than in the language of active processes. The relativity of judgement has been stressed, but the observer's reaction seems to be merely a prodigious compilation of weighted averages, nearly as quick as a computer.

In a lifted-weight experiment, Brown (1953) demonstrated that the way an object is identified or classified is critical. A relatively heavy tray, introduced during a series of judgements, had no regular effect upon the categories, although an object of exactly the same weight and similar to those in the series did have a marked effect when introduced at the same times as the tray. 'The anchor, to be effective, must be

perceived as a member of the same class of objects as the other weights and as having the attribute being judged by *O*. One factor . . . which influences the manner in which the stimulus-object is perceived is the dimension of similarity . . . ' (p. 210).

The anchor introduced in Brown's experiment had less of an influence when it was not judged directly than when it was categorized in the same way as the weights in the series. The effect of this variable and that of object similarity can both be attributed to a single variable, conceived relevance. For the requirements of the task a tray was less relevant than a weight; and when no weight judgement was required the anchor weight was isolated from the others in the series—to a degree or for some subjects it was irrelevant.

What are the essentials of an experiment to verify that in such tasks there is selectivity based on relevance, conceived as an activity of the observer rather than a direct function of the stimuli? The objects of the series must vary in two or more dimensions or characteristics, at least two of which are judged. After adaptation levels have been established a new object (anchor) is introduced, under conditions in which it may or may not be considered relevant to the series being categorized. A highly similar anchor is used, and it is not judged, but there must be evidence that it is observed. All subjects are exposed to the same objects and anchor in the same order and the same place, while a difference in relevance is induced by instructions, or by a secondary task which introduces no changes in the visual field.

There are three bases for predicting what should occur, each with different consequences: (1) If category judgements are completely determined by the magnitudes of all similar objects observed and by the background, when an anchor is introduced there should be a shift for each of the dimensions judged, the degree of shift depending upon the magnitude of the particular attribute of the anchor and the number of times it has been presented. (2) If subjects with one 'set' tend to conceive one dimension of the anchor as relevant, while those with another are more cognizant of the other dimension, the same anchor should produce different shifts for the two groups. (3) Finally, if judgements are influenced only by items identified as within the series, a non-judged object should have no effect on any subjective scales, other than that attributable to a small change in background. The second of these statements is the primary hypothesis of this study.

It is hypothesized that during a series of multi-dimensional judgements the repeated observation of a similar object of an extreme magnitude does not necessarily influence an observer's relative scales of judgement; presumably when it has no effect it has not been conceived as relevant. One attribute of the object can lead to a shift in a category scale, while another of the same object has little or no effect. This is not determined by the physical characteristics of the object, for in a different observer the results can be reversed. The relevance of a dimension can be experimentally manipulated and predicted by altering instructions and supplementary tasks.

## II. METHOD

*Subjects.* The subjects were twenty-eight British Naval Ratings who were participating concurrently in a number of other experiments at the Applied Psychology Research Unit. The four to six subjects available for a given 2-week period were always in the same experimental group. Typically their manner of judgement was quick rather than contemplative.

*Apparatus.* The stimuli were fifty thin, rectangular blocks, five each of ten different heights in equal intervals from 5 to  $12\frac{1}{2}$  in. There were ten widths, five of each width from 2 to 5 in. Moreover, each block was covered with one of ten selected shades of grey paper, from near white to near black. The ten shades of grey were selected from thirty produced by the Mitchell Colour Card Co., London, England. Their percentage reflectances were: 4.8, 6.5, 9.2, 20, 26, 33, 42, 48, 58 and 78. The combinations of height, width and lightness were systematically ordered so that there was no correlation between them. Prior to presentation the blocks were kept upright and in order in five trays, shielded from the subject's view. In a tray each height, width, and lightness was represented once. The order within the tray was randomly determined, but was the same for each subject. The blocks were presented on a sloping, four-foot-square, light-yellow background, each block dropping from a slot at the top of the background and coming to rest on a special catch below and to the right of the centre. The rotating catch that held the block was connected to a lever in front of subject, so that he could drop it at the end of the trial either into a bin on the left or one on the right. A magnet was positioned on the back of the board to hold a standard block in place during the second part of the experiment. The standard was  $12 \times 4\frac{1}{8}$  in. with a light surface (reflectance: 71%). The standard was held so that the bottom of it was horizontally aligned with the top of the comparison stimuli. The standard and comparison were at least 2 in. apart, with the centre of the board between. An interval timer operated a buzzer during the paced trials, warning subject of each new trial, and it operated a light indicating to *E* when the next block was to be released. A chin rest was used to maintain a constant viewing distance of 24 in.

*Procedure.* Subject was instructed that this was a test 'to learn how quickly and how well men can be trained to identify unfamiliar targets'. He was asked to call out whether each block was very tall, tall, medium height, short or very short; and whether it was very light, light, medium grey, dark or very dark. The order of the variables in the instructions indicated which judgement was to come first, and subject was to judge in that order. Half the subjects judged height first, and half lightness first. During the instructions one block was presented when the warning buzzer was demonstrated, and two additional blocks were shown while *E* repeated the possible categories of judgement twice more. Subject was not asked to respond on these four trials, although sometimes responses were volunteered. The differences in height, width and lightness of the first four blocks provided an introduction to the nature of variation to be expected. The repetitions of the categories were sufficient for most subjects to learn them, with only an occasional question or prompting during the practice trials that followed. Subject was told to take as long as he needed to make the judgements during the first session, and the next block was presented as soon as he had disposed of the preceding one. Three trays (thirty blocks) were used in this session, during which categories and the order of response were established, and subject was introduced to the full range of stimulus differences. The responses for this session were recorded, but as practice to establish consistency they were not included in the main analysis. At the second session subject was told that the blocks would come at regular intervals (11 sec), and was assured that he would still have plenty of time to respond. Six sessions followed on different days, with five trays presented on each day in a different sequence. Three sessions were conducted under the first condition, to establish an adaptation level and scale of judgement.

Then a new task was introduced, to accompany and precede the familiar one on each trial. Subject was instructed to compare each block to the standard before judging height and lightness, and if it were taller (lighter) than the standard to say 'reject', to make his two category judgements as before, and to dispose of it in the *left* bin. If it were the same height (lightness) or less (darker), to say 'accept', call out its height and lightness, and dispose of it in the *right* bin. For half the subjects height was the basis for the accept-reject decision and for half lightness, with order of judgement balanced. Three series, each on a different day, were given with the standard present. All subjects were exposed to exactly the same stimuli, in the same order, under uniform conditions of lighting, background and pacing. Only the basis of relevance for the supplementary task was varied. No direct judgement of the standard block was required.

*Scoring.* The frequency with which subject responded with each of the five judgement categories for each of the stimulus values was determined. There were fifteen height judgements for each value in the initial part and fifteen in the second, and the same number for lightness. The subject's shift score for a stimulus value was obtained by subtracting each category of the cumulative frequency distribution after the standard was introduced from that of the initial distribution, and selecting the maximum absolute difference as the score (Saltzinger, 1956).



In the example illustrated (Table 1), the shift score is  $-4$ . The negative value indicates a shift away from the tall standard, an increase in the frequency of a lower category of response.

Of the ten stimulus values presented during the experiment, only the eight intermediate ones were used in the analysis of the data. Subjects who had consistently classified the lowest value as very short or very dark could not possibly shift in the predicted direction. Similarly, a subject who initially always applied the highest category to the stimulus at the top of the scale could not shift upward, away from the standard as might be predicted.

Table 1. *Enumeration and cumulative frequencies of a subject's judgements of blocks of a given height, to illustrate shift*

	Judgements									
	Very short		Short		Medium		Tall		Very tall	
	F.	Cum. f.	F.	Cum. f.	F.	Cum. f.	F.	Cum. f.	F.	Cum. f.
Series I	0	0	0	0	1	1	9	10	5	15
Series II	0	0	0	0	0	0	14	14	1	15
Difference in cum. f.		0		0		+1		-4		0

F., frequency; Cum. f., cumulative frequency.

### III. RESULTS

The subjects for whom the height of the standard (anchor) was made relevant in the secondary task demonstrated a greater shift in category judgements of height than did the subjects for whom it was not (Fig. 1). The shift was negative, indicating an increase in adaptation level. The differences between groups in total shift scores, matched for heights, were significant ( $P < 0.02$ ) using Wilcoxon's test for paired replicates. The differences in shift for the two taller stimuli were small, but for those more removed from the anchor—10 in. and shorter—there were consistent differences between groups in the effects of the anchor upon scales of height judgement. Eight of the fourteen subjects in the relevant group demonstrated consistent shifts for all or nearly all of the heights, while only two of the fourteen in the less-relevant group give similar shifts. The criterion of consistency for this classification was at least six negative shift scores for the different heights, and not more than one positive one.

For lightness judgements the difference between groups in shift scores was also significant ( $P < 0.01$ ). As indicated in Fig. 2, there is evidence that even the less-relevant group shifted its categories of judgement for stimuli near the standard, so that degree of relevance influenced the extent of the effect rather than simply its presence or absence. Seven subjects in the relevant group displayed consistent shifts in lightness categories, applying the same criterion as for height, while four shifted in the less-relevant group.

Considering both groups of subjects, there were fourteen who revealed consistent shifts in judgement on one judged variable and not on the other, twelve of whom shifted as had been predicted from the relevance of the secondary task. Five of those in the height-relevant groups shifted consistently only in their height judgements, and seven in the lightness-relevant group shifted only in their lightness judgements. There were, in addition, three subjects who shifted consistently in judgements of both variables.

The groups of subjects were not random samples, and differences between groups of four to six subjects who participated at one time were confounded with treatment differences. For this reason subjects shift scores have been presented as basic data, rather than the sums of the frequencies of each category of response.

In the accept-reject task the differences in height and lightness between the standard and the stimuli within the series were discriminated accurately more than 90 % of the time, and there was no significant association between subjects' errors and shift scores.

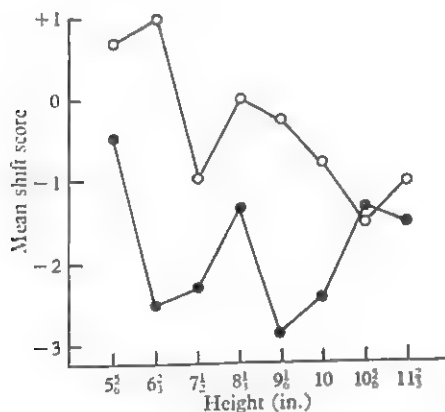


Fig. 1

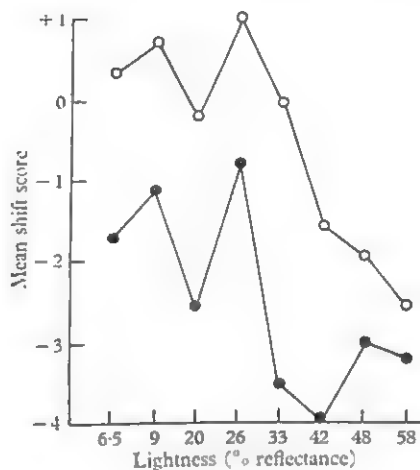


Fig. 2

Fig. 1. Shifts in height judgements. Mean shift scores are indicated for each group of fourteen subjects for blocks of the given heights. Values for the group for whom height was relevant in accepting or rejecting each block are designated 'relevant', and the height shift scores for the group for whom lightness was the basis for each decision are designated 'less relevant'. The height of the standard was 12 in. Shifts for the extreme height values (5 and 12 in.) are not included. ○—○ less relevant; ●—● relevant.

Fig. 2. Shifts in lightness judgements. Mean shift scores are indicated for each group of fourteen subjects for blocks of the given reflectance values. Means for the group for whom lightness was relevant in the secondary task for accepting or rejecting each block are designated 'relevant', and the lightness shift scores for the group for whom height was the basis for each decision are labelled 'less relevant'. The standard had a reflectance value of 71 %. Shifts for the extreme lightness values (4.8 and 78 %) are not included. ○—○ less relevant; ●—● relevant.

No general effect of the order in which subjects judged height and lightness was reliably demonstrated, but for the two lightest rectangles of the series there were much greater shifts when lightness was judged first, immediately after the comparison with the standard. Because of the high variability both in initial categories of response and in shift scores, further evidence is required to distinguish the effects of sequence in multi-dimensional judgement.

#### IV. DISCUSSION

The evidence is clear that persons' scales of judgement change and that these changes are predictable. Two simplified accounts of the processes by which category scales are modified have been distinguished by Stevens (1958), and neither is sufficient. One refers to a direct modification of a sensory system with each of a class of stimulus

events, a cumulative *adaptation* which automatically changes the perceived relationships between subsequent similar stimuli. The other attributes the change to mere semantic variation, to mercurial changes in arbitrary labels, not essentially related to perception or to underlying 'psychological continua'.

In the present experiment, as in certain studies by Helson and associates (Helson, Blake, Mouton & Olmstead, 1956; Helson, Dworkin & Michels, 1956), the shifts in the scale of judgements cannot be attributed to a specifiable change in physiological state or level of functioning which has been a direct function of excitation. The fact that the same standard against the same background produced only a shift in height judgements in one subject and only a shift in lightness judgements in another cannot readily be explained in terms of electrochemical states of the visual mechanism.

To recognize that conceptualizing is involved, however, that there is a classification or coding process, is not to deny that such changes are relevant to fundamental issues in the study of perception.

It is proposed that in perceiving and identifying the same object on different occasions its many attributes are not equally prominent. As context and the concept-structure of the individual change the subject's relative responsiveness to the different characteristics of the object is modified. Figuratively speaking, in one context the height of the object may be one of its most distinct characteristics, and in another it may be almost lost among hundreds of qualities. The object is perceived differently. This is more than a mere change in tag.

In the present experiment by changing the context—the relevance of the anchor to the judgemental task—a difference was induced in the way the anchor was perceived by some subjects, a difference reflected in modified categories of response. Although not itself judged, for one group of subjects not only was the anchor conceived as a member of the series, but it was perceived with its height salient, just as the height of each block in the series was salient for many members of this group. For those in the other group whose lightness judgements shifted, lightness was a more prominent characteristic of both the anchor and the other stimuli.

A contrast or repulsion effect (Helson & Nash, 1960) was produced by the anchor, rather than assimilation as reported by Sherif, Taub & Hovland (1958) for an anchor at the end of the scale. In the present experiment the anchor was one step from the end, but this is not believed to be the crucial difference. Although it was a standard for the accept-reject decision, the block placed on the board in Part II was not designated a standard for categorizing and classified by *E* as very tall or very light. Had it been so designated, not only would more shifts probably have occurred, but many more would have been in the opposite direction. Rather, for the subjects that shifted, the obtained effect was more like that of an added object in the series of like magnitude presented many times. The direction of the effect and its decrease as a function of distance from the anchor were in accord with adaptation level theory.

## V. SUMMARY

To verify a selective process in the evolution of an observer's category scales of judgement, all subjects repeatedly judged both the height and lightness of each of fifty rectangles under two conditions. In the first, all of the 28 subjects established



adaptation levels and consistent category responses for each of the variables under identical conditions; while in the second part of the experiment a standard (anchor) was introduced and each group of fourteen subjects was given a different supplementary task. One group was instructed to compare each block to the standard and to accept or reject it on the basis of its height, as well as to categorize the stimuli as before. For the other group the accept-reject decisions were based on the lightness of the standard. The supplementary task tended to make one characteristic of the standard salient more relevant, without producing a change in the visual field.

It was demonstrated that repeatedly observing a similar object of extreme magnitude does not necessarily produce a shift in the judgemental scale, for many subjects categories did not shift significantly in the second part of the experiment. There were significant differences in shift scores between the two groups. Subjects for whom height was more relevant tended to shift height categories but not those for lightness, and those subjects for whom the lightness of the anchor was more relevant tended to shift only in lightness judgements.

Whether or not an observed object had an effect upon the individual's scale of judgement, and which of its attributes was effective, were both a function of its relevance within the task.

The study was conducted while the author was a Visiting Fellow at the Applied Psychology Research Unit, Medical Research Council, Cambridge, England, during the tenure of a Special Research Fellowship granted by the National Institute of Mental Health, U.S. Public Health Service. The assistance of Mr Donald E. Broadbent, who made this study possible, is gratefully acknowledged. Thanks are due to Mr C. J. Deane who constructed the apparatus.

#### REFERENCES

- BALL, J. H. (1953). The influence of background and residual stimuli upon the measurement of attitudes. Unpublished master's thesis, University of Texas.
- BROWN, D. R. (1953). Stimulus similarity and the anchoring of subjective scales. *Amer. J. Psychol.* **66**, 199-214.
- HELSON, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *Amer. J. Psychol.* **60**, 1-29.
- HELSON, H., BLAKE, R. R., MOUTON, JANE S. & OLMSTEAD, J. A. (1956). The expression of attitudes as adjustments to stimulus, background and residual factors. *J. Abnorm. Soc. Psychol.* **52**, 314-22.
- HELSON, H., DWORKIN, R. S. & MICHELS, W. C. (1956). Quantitative denotations of common terms as a function of background. *Amer. J. Psychol.* **69**, 194-208.
- HELSON, H. & NASH, M. C. (1960). Anchor, contrast, and paradoxical distance effects. *J. Exp. Psychol.* **59**, 113-21.
- MICHELS, W. C. & HELSON, H. (1949). A reformulation of the Fechner law in terms of adaptation-level applied to rating-scale data. *Amer. J. Psychol.* **62**, 355-68.
- MICHELS, W. C. (1953). An interpretation of the bril scale of subjective brightness. *J. Opt. Soc. Amer.* **44**, 70-74.
- MICHELS, W. C. & DOSER, B. T. (1955). Rating scale method for comparative loudness measurements. *J. acoust. Soc. Amer.* **27**, 1173-80.
- NASH, MYRTLE C. (1950). A quantitative study of effects of past experience on adaptation-level. Unpublished doctoral dissertation. Bryn Mawr College.

- SALZINGER, K. (1956). Techniques for computing shift in a scale of absolute judgment. *Psychol. Bull.* **53**, 394-401.
- SHERIF, M., TAUB, D. & HOVLAND, C. L. (1958). Assimilation and contrast effects of anchoring stimuli on judgments. *J. Exp. Psychol.* **55**, 150-5.
- STEVENS, S. S. (1958). Adaptation level vs. the relativity of judgment. *Amer. J. Psychol.* **71**, 663-46.

*(Manuscript received 17 October 1961)*

## COGNITIVE CONTROLS OF ATTENTION AND INHIBITION: A STUDY OF INDIVIDUAL CONSISTENCIES\*

BY RILEY W. GARDNER AND ROBERT I. LONG

*The Menninger Foundation*

Results of a previous study (Gardner, 1961) suggested that independent principles of cognitive control are involved in (a) extensiveness and (b) selectiveness of attention deployment. The present study extends these results by demonstrating that differential predictions can be made even when criterion scores for the two control principles are obtained from a single set of judgements. It is also shown that individual differences in selective attention can be observed whether compelling irrelevant stimuli enclose, are enclosed by, or are interpolated among relevant stimuli, and that two forms of inhibition—extensiveness of scanning and inhibition of irrelevant motoric responses—may be associated.

### I. INTRODUCTION

Gardner (1961) recently demonstrated that two independent dimensions of individual consistency in attentional behaviour are related in predictably different ways to the experiencing of two types of illusions. One of these dimensions was originally conceived of as a cognitive control principle concerning extensiveness of scanning (Gardner, Holzman, Klein, Linton & Spence, 1959). Gardner (1961) used apparent magnitude scores for judgements of plain circular disks to represent this dimension. Following Piaget, whose extensive studies of relations between attention deployment and apparent magnitude are summarized in a recent monograph (Piaget, 1961), Gardner assumed that these apparent magnitude scores were correlates of extensiveness of scanning, with the extensive scanner showing a consistent counter-active approach resulting in minimal 'error of the standard', i.e. minimal overestimation of the constant stimulus, which Piaget attributes to the repetition of centrations upon it during judgement. (In Gardner's study, as in the present one, each look at the standard was operationally defined as one 'centration'.) Gardner & Long (1962) recently confirmed the hypothesized negative relationship between extensiveness of scanning (e.g. in the form of number of centrations on standard and comparison stimuli) and the apparent magnitude of these disks *per se*. Their findings made possible a more precise evaluation of relations between field articulation and extensiveness of scanning.

Gardner & Long (1962) also found that extensiveness of scanning in size estimation tests was related to ratings of 'generalized delay' in the psychoanalytic sense, which is not limited to temporal delay but includes all cognitive transformations or patterned channellings of impulse expression (each of which implies inhibition of immediate motoric expression of impulses). These ratings were based on Rorschach Test protocols. Extensiveness of scanning was also related to duration of scanning protocols of inkblots before production of first responses. In addition, the Rorschach protocols of extensive scanners were rated higher than those of limited scanners in indications of

\* This investigation was supported by grant M-2454 from the National Institute of Mental Health, United States Public Health Service. We are indebted to Dr Gardner Murphy for a helpful reading of a draft of this paper.



either of two defence mechanisms, each of which has special implications for the inhibition of impulse expression: isolation, apparently leading to obsessive indecision, and projection, apparently leading to suspicious cautiousness. The two defence ratings were not significantly associated. In another recent study of 'generalized delay', Spivack, Levine & Sprigle (1959) showed that rapid colour-naming in the Stroop colour-word test is negatively associated with time estimations of 30 and 60 sec. That is, persons who effectively inhibit verbalization of the inappropriate colour words, and thus name the colours faster, also show relatively great inhibition in estimating the lengths of short periods of time. Colour-word and time estimation performances were also significantly related to I.Q. scores. Although further exploration of the relations to particular subtest scores of the intelligence test is required for understanding these results, their study suggests the generality of a control principle relevant to inhibition of impulse expression. Holt (1960) recently provided evidence that subjects with disrupted colour-word test performances show relatively great difficulty in controlling impulses and relative inability to use their primary-process thinking in effective ways. Klein (1954) is among those who have shown that subjects who can inhibit irrelevant motoric responses in the colour-word test are also those who can inhibit or counteract the effect of a need upon cognition. He spoke of colour-word test performances in terms of a dimension of constricted and flexible control, the former leading to disruption of performance in the colour-word test, the latter leading to effective inhibition of the irrelevant motoric responses. The recent studies referred to here seem to support his view. The present study provided an opportunity to test the hypothesis that reading time in a colour-word test is also related to inhibition of response in the form of extensiveness of scanning.

A second dimension of individual consistencies, conceived of by Gardner *et al.* (1959) as a cognitive control principle concerning field articulation, is apparent in the articulation of experience in a wide variety of situations originally referred to by Witkin, Lewis, Hertzman, Machover, Meissner & Wapner (1954) as sampling 'field dependence'. In embedded figures tests and related laboratory tests that provide criterion scores of this dimension, articulation of the field apparently takes the form of selective attention to (i.e. concentration upon) relevant *vs.* irrelevant stimuli (Gardner *et al.* 1959). This interpretation recently allowed prediction of responses to a cluster of intellectual ability tests previously assumed to assess different abilities (Gardner, Jackson & Messick, 1960) and recall of two lists of highly similar words (Gardner & Long, 1961), in addition to the experiencing of one type of illusion (Gardner, 1961).\*

Prediction of selectiveness of attention to relevant *vs.* irrelevant parts of three types of stimulus fields was included in the present study because of the discrepancy between Witkin's statements that embedded figures and related tests assess the capacity to extract embedded items from their *surrounds* or *contexts* (see, e.g. Witkin *et al.* 1954; Witkin, 1959; Witkin, Karp & Goodenough, 1959) and results obtained by others (e.g. Jackson, 1955; Longenecker, 1956; Podell & Phillips, 1959; Gardner *et al.* 1959) suggesting that these individual differences obtain whether the compelling

\* Field-articulation scores are related only to certain intellectual ability scores and certain subtest scores of intelligence scales (see Witkin, 1959). To 'partial out' I.Q. in the present study therefore seemed inappropriate. The general issue of relations between cognitive controls and 'intelligence' is discussed in detail by Gardner *et al.* (1959).

irrelevant stimuli comprise the surround; an embedded item; or distractors lacking cohesive organization. The concentric-circles illusion of Delbœuf was used to assess selectiveness of attention to a relevant stimulus enclosed by a compelling, illusion-producing irrelevant stimulus.\* Piaget (1961) explains this illusion in terms of his 'law of relative centrations'. We would add only that, under appropriate stimulus conditions, individual differences in selectiveness of attention lead to predictable variations in the phenomena subsumed under this law. Since the illusion is much greater for the enclosed circle (enhanced magnitude) than for the enclosing circle (reduced magnitude), a special figure was included in the study to assess selectiveness of attention to the enclosing portion of a stimulus field when the enclosed portion is compelling but irrelevant to the judgement required by the experimenter. Recording of gross eye movements during judgements of these two illusion-producing figures provided a stringent test of the independence of selectiveness of attention and extensiveness of scanning. The authors also constructed a new word-reading test to show that the control principle is involved in performance when compelling irrelevant stimuli are simply interpolated among relevant stimuli.

## II. HYPOTHESES

The following hypotheses were tested. (1) Individuals differ consistently in extensiveness of scanning (number of centrations on standard stimuli) in making size judgements of varied stimuli. (2) Extensiveness of scanning and selectiveness of attention, the latter as indicated by mean solution time in an embedded figures test, are independent (i.e. uncorrelated) aspects of attentional behaviour. (3) Inhibition in the form of extensiveness of scanning before commitment to a size judgement is related to inhibition of irrelevant motoric responses in a colour-word test (i.e. is negatively related to colour-word interference score). (4) Extensiveness of scanning is negatively related to the apparent magnitude of a circular disk judged by means of a variable circle of light. (5) Mean solution time in the embedded figures test is positively related to illusion effect in size judgements of either enclosed or enclosing portions of stimulus fields. (6) Mean solution time in the embedded figures test is positively related to interference (i.e. increment in reading time) in a word-reading test in which irrelevant letters are interpolated among relevant letters forming words.

## III. METHOD

### (i) *Subjects*

The sixty female subjects were housewives, college and business school students. The age range was 16-43, mean age, 23.0.

### (ii) *Procedures*

#### (a) *Size estimation tests*

The apparatus used to produce the variable circle of light was that used by Gardner (1961) and by Gardner & Long (1962) for Test I of their group of size estimation

\* It was assumed that a related hypothesis involving the Gottschaldt test and the Delbœuf illusion was not confirmed by Rabe (1955) because a small (10 mm.) enclosed circle was used and because no control was provided for errors of the standard (see Piaget, 1961).

procedures. This circle of light appeared  $\frac{1}{2}$  in. above and  $6\frac{3}{4}$  in. to the right of each standard stimulus. Subjects made one ascending followed by one descending judgement of each standard from starting points 10 mm. below and above the size of the standard.

In the experimental condition, the enclosed standard was a 35 mm. circle surrounded by a 50 mm. circle. The enclosing standard was a 50 mm. circle containing drawings suggesting the faces and/or heads of a multitude of persons, behind which short vertical lines suggested a back-drop. Each of these figures was drawn in India ink on a  $4\frac{1}{2}$  in. (horizontal)  $\times$  9 in. (vertical) white cardboard background.

In the control condition which followed, subjects made judgements of 35 and 50 mm. circles identical to those described above except that the enclosing or enclosed portions were absent.

Thirty-nine of the subjects also made judgements of a black disk 50 mm. in diameter and 6 mm. thick to the back of which a smaller metal disk was attached. This disk was affixed to a  $4\frac{1}{2}$  in. (horizontal)  $\times$  9 in. (vertical) flesh-coloured wooden background containing an embedded magnet.

Apparent magnitude scores were computed for each of the five standard stimuli. The mean difference between apparent magnitude in the experimental and control conditions was 2.46 mm. for the 35 mm. circle and 0.63 for the 50 mm. circle. Pearson  $r$ 's between apparent magnitude scores for the control and experimental conditions of the two circle sizes were 0.58,  $P < 0.001$ , for the 35 mm. circle and 0.48,  $P < 0.001$ , for the 50 mm. circle. Residual scores were therefore used to represent the illusion effects. These consisted of the difference between the subject's actual score and her predicted score under each illusion-producing condition, the prediction being based upon the regression of the actual score on the appropriate control score. These residual illusion-effect scores, correlated 0.47,  $P < 0.001$ , were also summed to provide a total illusion-effect score.

The experimenter observed and manually recorded the number of centrations on each of the five standard stimuli.

#### (b) *Embedded figures test*

The first twelve items of Witkin's (1950) embedded figures test were administered according to Witkin's standard procedure except that the time limit was 3 min., rather than 5 min. The score used was the mean time required to identify the twelve simple figures correctly.

#### (c) *Colour-word test*

This test was administered and scored as described by Gardner *et al.* (1959). Mean reading (i.e. colour-naming) time for colour strips was 57.1 sec., for colours printed in incongruous colour words, 98.9 sec. The  $r$  between these scores was 0.64,  $P < 0.001$ . The residual interference score for each subject consisted of the difference between her reading time in the colour-word portion of the test and her predicted reading time, the prediction being based on the regression of reading time for the colour-word combinations on reading time for the colour strips.

#### (d) *Word-reading test*

The word-reading test was administered to forty-nine of the subjects. In the control portion of the test, given first, the subject was required to read a page of twenty-eight



three-letter words as rapidly and accurately as possible. These stimuli, typed in black capital letters on white paper, appeared in seven lines of four words. One blank space appeared between the letters of these words. In the experimental condition, given after a 1 min. pause, the subject read a second page of twenty-eight three-letter words. On this page, the spaces between the relevant letters were filled with irrelevant letters.

Mean reading time for the control condition was 13.9 sec., for the experimental condition, 67.0 sec. Since the  $r$  between reading time in the control and experimental conditions was 0.30,  $P < 0.05$ , a residual interference score was computed to represent the increment in reading time attributable to the presence of the interpolated letters.

#### IV. RESULTS

Hypothesis 1, that individuals are consistently different in extensiveness of scanning, is confirmed by the highly significant  $r$ 's in Table 1. These results are similar to those obtained by Gardner & Long (1962) for a more inclusive syndrome of scanning behaviours including number of centrations on the standard.

Table 1. *Intercorrelations among number of centrations scores†*

	1	2	3	4
1. 35 mm. circle	—	—	—	—
2. 50 mm. circle	0.71***	—	—	—
3. 35 mm. circle enclosed by 50 mm. circle	0.63***	0.78***	—	—
4. 50 mm. circle enclosing drawing of persons	0.46***	0.65***	0.79***	—
5. 50 mm. black disk	0.62***	0.87***	0.81***	0.68***

†  $N = 60$ , except for the black disk, for which  $N = 39$ .

\*\*\*  $P < 0.001$ .

Hypothesis 2, that extensiveness of scanning and selectiveness of attention are uncorrelated, is confirmed by the non-significant  $r$ 's ranging from  $-0.20$  to  $0.02$  between mean solution time in the embedded figures test and the five extensiveness of scanning scores.

Hypothesis 3, that extensiveness of scanning and colour-word interference are negatively associated, is at least partially confirmed by the fact that  $r$ 's between the five extensiveness of scanning scores and the interference score range from  $-0.22$  to  $-0.38$ , all in the predicted direction, with two significant at  $P < 0.05$ .

Results pertinent to Hypotheses 4 and 5 are presented in Tables 2 and 3. It is apparent from Table 2 that the predicted negative relationship of extensiveness of scanning to apparent magnitude of the 50 mm. black disk is confirmed. That extensive scanners were not simply more accurate in judgements of the black disk is indicated by the fact that mean constant error with this figure was  $-0.88$ . The predicted positive relationship of the embedded figures score to illusion effects with both enclosing and enclosed circles is also confirmed.

It is also apparent from Table 2 that both extensiveness of scanning and selectiveness of attention may be determinants of the apparent magnitude of the 35 and 50 mm. control circles. Although the former might have been anticipated, the latter was not. Further consideration of these figures and their backgrounds seemed to make this unanticipated result easily understandable. In contrast to the black disk, which is highly impressive against a relatively weak flesh-coloured background, both

Table 2. *Pearson correlations of scanning and embedded figures scores with apparent magnitude scores†*

	Scanning: number of centrations on standard	Embedded figures: mean solution time
Apparent magnitude, 50 mm. black disk	-0.49**	0.16
Apparent magnitude, 35 mm. circle	-0.29*	0.30*
Apparent magnitude, 50 mm. circle	-0.20	0.26*
Apparent magnitude, enclosed 35 mm. circle	0.04	0.39**
Apparent magnitude, enclosing 50 mm. circle	-0.06	0.47***
Illusion effect, enclosed 35 mm. circle	0.18	0.29*
Illusion effect, enclosing 50 mm. circle	-0.02	0.38**
Illusion effect, total	0.14	0.46***

†  $N = 60$ , except for the black disk, for which  $N = 39$ .\*  $P < 0.05$ .\*\*  $P < 0.01$ .\*\*\*  $P < 0.001$ .Table 3. *Pearson correlations among apparent magnitude scores†*

	1	2	3	4
1. 35 mm. circle	—	—	—	—
2. 50 mm. circle	0.74***	—	—	—
3. 35 mm. circle enclosed by 50 mm. circle	0.58***	0.61***	—	—
4. 50 mm. circle enclosing faces	0.55***	0.48***	0.42***	—
5. 50 mm. black disk	0.44**	0.40*	0.25	0.25

†  $N = 60$ , except for the black disk, for which  $N = 39$ .\*  $P < 0.05$ .\*\*  $P < 0.01$ .\*\*\*  $P < 0.001$ .

these circles and their white backgrounds have sharply defined contours. Presentation of the circles against  $4\frac{1}{2}$  in. wide backgrounds, compared to the variable light circle against a  $9 \times 9$  in. background, could have produced contrast effects leading to illusions of enhanced magnitude. To overcome these illusions would presumably require that attention be directed selectively to the standard and comparison stimuli and withheld from the irrelevant backgrounds. A similar relationship, between solution time in the embedded figures test and illusion effect in equating lines bounded by large and small squares, was in fact obtained by Gardner (1961) in his study of illusions. The intercorrelations of apparent magnitude scores presented in Table 3 seem to support the inference that both extensiveness of scanning and selectiveness of attention were involved in the control judgements, whereas selectiveness of attention predominated in the experimental judgements. Thus, all the  $r$ 's are significant *except* those between the apparent magnitude of the enclosed and enclosing circles and that of the black disk. The unintended requirement to attend selectively in the control judgements obviously reduced the likelihood of confirming the predicted relationship of the embedded figures score to the two residual illusion-effect scores.

Hypothesis 6, that mean solution time in the embedded figures test is related to reading-speed increment in the presence of interpolated irrelevant stimuli, is also confirmed. The  $r$  between the embedded figures test score and the interference score for the word-reading test is 0.36,  $P < 0.01$ .

## V. DISCUSSION

Gardner *et al.* (1959) have stressed that cognitive controls are enduring patterns, strategies, or 'programmes' of cognitive behaviour. In the psychoanalytic ego-psychological framework in which they were conceived, controls are viewed as enduring cognitive structures that, like defence mechanisms, presumably emerge in the course of development from the interaction of genetic and experiential determinants. These structural arrangements of secondary-process functions are assumed to shape the expression of particular adaptive intentions under particular classes of conditions. The present results support this general assumption and point to the value of further studies based on differential predictions for pairs or small groups of control principles. Thus, extensiveness of scanning has predictable consequences for performance under certain conditions, field articulation under related but distinctly different conditions. Our results also indicate clearly that expression of the field-articulation control principle in situations involving relevant and irrelevant stimuli is not limited to the extraction of items from embedding contexts. Rather, the degree of field-articulation is a general characteristic of the individual's cognitive behaviour requiring no special assumptions about the *organizational* properties of the 'field'. It is noteworthy, however, that these individual consistencies in selectiveness of attention can be observed only when the irrelevant stimuli, ideas, or memories (Gardner & Long, 1961) are quite compelling.

Recent studies of the interaction of cortical and reticular functions—as these are involved in response to selected aspects of incoming stimulation—may elucidate some of the specific mechanisms involved in this control principle (see, e.g. Vernon, 1959; Magoun, 1958; Jasper, 1958).

Taken together with the earlier findings of Spivack *et al.* (1959), Holt (1960), Klein (1954), and others, our finding that inhibition in the form of extensiveness of scanning may be related to inhibition of irrelevant motoric responses in the colour-word test (which could involve both attentional and motoric functions) calls for further exploration. The diversity of forms of inhibition that appear to be linked in the present study and the related earlier studies may reflect a general control principle concerning inhibition that has implications for the mediation of motive expression in a variety of situations.

## REFERENCES

- GARDNER, R. W. (1961). Cognitive controls of attention deployment as determinants of visual illusions. *J. Abnorm. Soc. Psychol.* **62**, 120-7.
- GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, H. B. & SPENCE, D. P. (1959). Cognitive control: A study of individual consistencies in cognitive behavior. *Psychol. Issues*, **1**, no. 4.
- GARDNER, R. W., JACKSON, D. N. & MESSICK, S. J. (1960). Personality organization in cognitive controls and intellectual abilities. *Psychol. Issues*, **2**, no. 4 (whole no. 8).
- GARDNER, R. W. & LONG, R. I. (1961). Field-articulation in recall. *Psychol. Rec.* **11**, 305-10.
- GARDNER, R. W. & LONG, R. I. (1962). Control, defence, and centration effect: A study of scanning behaviour. *Brit. J. Psychol.* **53**, 129-40.
- HOLT, R. R. (1960). Cognitive controls and primary processes. *J. Psychol. Res.* **4** (3), 1-8. (Unpublished Ph.D. thesis) Purdue University.
- JACKSON, D. N. (1955). Stability in resistance to field forces. (Unpublished Ph.D. thesis) University of Wisconsin.
- JASPER, H. (1958). Reticular-cortical systems and theories of the integrative action of the brain. In *Biological and Biochemical Bases of Behavior*, pp. 37-61. H. F. Harlow & C. N. Woolsey. Madison: University of Wisconsin Press.



- KLEIN, G. S. (1954). Need and regulation. In *Nebraska Symposium on Motivation*, pp. 224-74. M. R. Jones (Ed.). Lincoln: University of Nebraska Press.
- LONGENECKER, E. D. (1956). Form perception as a function of anxiety, motivation, and the testing situation. (Unpublished Ph.D. thesis) University of Texas.
- MAGOON, H. W. (1958). Non-specific brain mechanisms. In *Biological and Biochemical Bases of Behavior*, pp. 25-36. H. F. Harlow & C. N. Woolsey. Madison: University of Wisconsin Press.
- PIAGET, J. (1961). *Les Mécanismes Perceptifs*. Paris: Presses Universitaires de France.
- PODELL, J. E. & PHILLIPS, L. (1959). A developmental analysis of cognition as observed in dimensions of Rorschach and objective test performance. *J. Personality*, **27**, 439-63.
- RABE, A. (1955). Individual differences in orientation in perceptual and cognitive tasks. *Canad. J. Psychol.* **9**, 149-54.
- SPIVACK, G., LEVINE, M. & SPRIGLE, H. (1959). Intelligence test performance and the delay function of the ego. *J. Consult. Psychol.* **23**, 428-31.
- STROOP, J. R. (1935). Studies of interference in serial verbal reaction. *J. Exp. Psychol.* **18**, 643-61.
- VERNON, M. D. (1959). Perception, attention and consciousness. *The Advancement of Science*, **62**, 111-23.
- WITKIN, H. A. (1950). Individual differences in ease of perception of embedded figures. *J. Personality*, **19**, 1-15.
- WITKIN, H. A. (1959). The perception of the upright. *Sci. Amer.* **200**, 50-6.
- WITKIN, H. A., KARP, S. A. & GOODENOUGH, D. R. (1959). Dependence in alcoholics. *Quart. J. Stud. Alc.* **20**, 493-504.
- WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B. & WAPNER, S. (1954). *Personality through Perception*. New York: Harper.

(Manuscript received 25 September 1961)

## SEQUENTIAL RECALL OF A MIXED LIST

BY ROBERT T. GREEN AND GRAHAM HARDING\*

*University College London*

A previously reported finding that an item with a high probability of recall will not necessarily be emitted any earlier in a recall series was submitted to closer investigation.

The anomalous result was reproduced, using the original experimental design. Two further groups were introduced to control context effects and item uniqueness. More careful analysis shows that the factors governing order of emission during recall may be distinguished from those affecting the overall probability of recall. The two processes are distinct and, in some circumstances, independent.

After eliminating various alternative hypotheses the outcome in the present instance was found to depend largely on previously learned associations of meaningful material. The implications of this with regard to the von Restorff effect are discussed.

During the course of this experiment an interesting side effect intruded itself on the data. For reasons not at present clear, there was a significant tendency for the recall of paralog to be disturbed in the presence of a noun, and, to a lesser extent, for the recall of nouns to be disturbed by the presence of a paralog.

### PROBLEM

In a recent paper by Saltz & Newman (1959) passing reference is made to an interesting phenomenon. The main purpose of their experiment was to expose the inadequacy of a serial interference interpretation of the von Restorff effect. As part of their design they presented a serial learning task to two groups of subjects and required free recall after one presentation of a list. These two groups were distinguished by the type of material presented. What was termed the Isolated group was given a list of thirteen items, twelve being common nouns and one a paralog, *gojey*, in the seventh serial position, i.e. a pronounceable nonsense item was isolated in a list of familiar meaningful items. The non-Isolated group was given the same list except for the seventh item, which was replaced by *office*. As was to be expected, the isolated item, *gojey*, was recalled significantly more often than the non-isolated *office*. Both less expected was the finding that, although *gojey* had a far higher probability of being recalled, it was not emitted any earlier in the recall series than *office*. Both items tended to appear around the fifth serial position during recall, despite the fact that over 80 % of the subjects learning the isolated list recalled *office*.

30 % of the subjects learning the non-isolated list recalled *office*. Saltz & Newman (1959, p. 449) state: 'Thus, despite the fact that the isolated word was recalled more frequently than the control word, isolation did not result in early emission of the isolated term. This result appears to be contrary to expectations from von Restorff's "figure-ground" theory and Green's (1956) "surprise" hypothesis. However, these theories have not been enunciated in sufficient detail to permit unequivocal predictions on this point.'

In point of fact, the result is so curious that it would be difficult to imagine any single factor theory accommodating itself to the data. Commonsense leads us to expect that if an item has a high probability of being recalled this means that it will

\* Now at the Industrial Psychology Laboratory, College of Advanced Technology, Birmingham  
Gen. Psych. 53, 4

tend to occur in the shorter as well as the longer recall lists, whereas an item with a low probability of recall would be expected to appear mainly in the longer lists recalled. This commonsense conclusion is implicit in any general theory of rote learning, such as that put forward by Hull *et al.* (1940). More limited theories, such as the extension of Marbe's law by Bousfield *et al.* (1956), make the same sort of prediction, although it must be remembered that these theories have been concerned, either tacitly or explicitly, with homogeneous material. With such material the commonsense expectation is borne out empirically, provided no other factor is permitted to influence the outcome. *Vide*, e.g. Deese & Kaufman (1957).

An analogy may help to make the paradox more obvious. Let us represent the items presented as sections of a roulette wheel. The size of the section representing each item corresponds to the degree to which it has been learned. Well-learned items take a larger section of the wheel and poorly learned items a smaller section. Recall is represented by spinning the wheel in the usual way, an item being selected for recall when the ball comes to rest in its sector. This sector is then eliminated and the next item is selected in the same way. Clearly, the better an item has been learned and the larger the sector representing it, the sooner it will be selected. If the process is cut short before all the items have been selected, those items poorly learned will tend to get left out. In other words, there must be a correlation between the probability of an item's being selected at all, and the serial position in which it is reproduced. The more probable its selection, the earlier will it be selected.

Any deviation from this rule indicates that the wheel is being influenced in a systematic manner. If, in a particular case, an item with a high probability of recall is selected no earlier than another item with a low probability of recall under the same circumstances, then some other factor must be counteracting the tendency inherent in the system. Quite evidently, no single factor theory can accommodate this anomalous type of result. Whatever accounts for the probability of an item's being recalled cannot supply a sufficient explanation for the serial position in which it is recalled.

This leaves us with the problem of pinpointing the hidden factor or factors operating to produce such an outcome. Several possibilities suggest themselves.

(1) The simplest explanation is that the average number of items recalled by the Isolated group was greater than for the non-Isolated group. In this case, although the absolute serial position of recall for the items does not differ, *gojey* might appear relatively early in terms of list-length recalled. Saltz & Newman do not give total recall scores, so this hypothesis must be held in reserve, although it is unlikely that they would overlook such an obvious explanation.

(2) Item uniqueness. Maybe there is something peculiar about either *office* or *gojey*. This seems unlikely, but no doubt an imaginative psychoanalyst could suggest some strange quality of *gojey* leading to its suppression to the last moment because of its threatening nature.

(3) There is the possibility that *gojey* exerts a disrupting influence, so that subjects become confused and emit a number of irrelevant nouns or other incorrect responses before getting round to *gojey* itself. Since we are not told anything about the incidence of errors during recall in the original experiment, this hypothesis is hard to evaluate as matters stand.



(4) Context effects. Saltz & Newman made their comparison by changing the critical item, but what happens if this item is held constant, and isolation and non-isolation produced by changing the contextual items? Commonsense tells us that if *office* were isolated amongst a list of paralogs it would be recalled both more frequently and earlier than when massed among a list of nouns, but Saltz & Newman's results make this prediction look a little less certain. Exactly the same argument holds for a comparison of *gojey* in the two contexts.

(5) Meaning and previously learned associations. Whereas *office* is a meaningful word and has only to be recognized as occurring in the list presented, *gojey* has to be learned also *as itself*. It is a nonsense item, without established associations in a familiar language system. Even its spelling has to be learned and remembered for this one special occasion.

#### PROCEDURE

Four groups were used, two corresponding to those of Saltz & Newman already referred to, and two others with paralogs as contextual material. The actual lists appear in Fig. 1.

Lists O-N and G-N		Lists O-P and G-P	
Serial positions			
1.	Dinner	1.	Sagrole
2.	Wagon	2.	Zumap
3.	Village	3.	Meardon
4.	Heaven	4.	Neglan
5.	Captain	5.	Polef
6.	Zebra	6.	Balap
7.	{ Office	7.	{ Office
	{ Gojey		{ Gojey
8.	Jewel	8.	Latok
9.	Garment	9.	Gokem
10.	Army	10.	Nares
11.	Insect	11.	Quipson
12.	Kitchen	12.	Tarop
13.	Money	13.	Volvap

Fig. 1. Four lists used. O-N and G-N refer respectively to *office* and *gojey* among nouns. O-P and G-P refer respectively to *office* and *gojey* among paralogs.

Each group contained twenty subjects, all University students. A small proportion were drawn from the Psychology Department, but these had less than 2 months of University life behind them and may therefore be fairly considered naïve. Age and sex differences were distributed evenly over the groups.

Instructions to subjects were as follows: 'This is a memory experiment. I shall show you a series of items, one at a time. The items will appear in this window'. (Points to memory drum window.) 'After you have had a chance to learn the material, you will have  $1\frac{1}{2}$  min. in which to write on the slip provided as many of the items as you can recall. The order makes no difference. I am interested only in the number of items you can recall correctly. Please attend carefully and avoid making any comments.'

Each item was displayed for 2 sec. The time allowed for recall proved ample for all subjects.

## RESULTS

These are set out in Table 1. The first two columns, O-N for *office* among nouns, and G-N for *gojey* among the same nouns may be compared directly with the corresponding two groups used by Saltz & Newman. It will be seen that their results are replicated with a fair degree of precision. The only difference worth noting is that the present results accentuate the paradox, since although *gojey* is recalled far more frequently than *office*, it is emitted later in the recall series.

Neither total number of items recalled, nor taking errors into account resolves the paradox. This effectively eliminates hypotheses 1 and 3.

Table 1. *Comparison of four groups*

	O-N	G-N	O-P	G-P
No. of subjects recalling critical item	5	18	19	5
Mean serial position of recall with errors	4.4	6.1	2.1	3.2
Mean serial position of recall without errors	4.4	6.1	1.5	2.0
Mean number of items recalled with errors	8.6	8.3	5.1	5.4
Mean number of items recalled without errors	8.6	8.1	1.8	3.0
Serial position of recall with errors (%)	51	73	40	59
Serial position of recall without errors (%)	51	75	80	67

The next interesting comparison is between groups O-N and O-P, where *office* is held constant and the contextual items changed. Here we find the result that common-sense would suggest likely. When isolated among paralogs, *office* is remembered far more frequently, and is also emitted earlier in terms of absolute serial position, whether errors are taken into account or not. This result would seem to favour hypothesis 4, until the relative or percentage serial positions are considered. Including errors, *office* is emitted 40 % of the way through the recalled list. If errors are taken out, *office* is emitted 80 % of the way through the list of correctly recalled items. This compares not altogether favourably with 51 % for *office* in a context of nouns, where no false responses were made.

The weakness of hypothesis 4 is revealed more clearly by making the parallel comparison between groups G-N and G-P. Again, the critical item, *gojey*, isolated among nouns, is recalled far more frequently than when massed among paralogs. In serial position terms, however, absolute or relative, with or without errors, *gojey* is recalled later when isolated than when massed. This just about disposes of the idea that Saltz & Newman had simply failed to make a relevant comparison by changing the critical item instead of the context.

This leaves us with hypotheses 2 and 5. A Freudian interpretation is always possible in principle, but in this case does not look too promising. An alternative *post hoc* explanation in a related category was indicated when one subject volunteered the information that he had withheld the isolated item to the end of his recall list, simply because he had learned it so well that he was in no danger of forgetting it. He therefore concentrated on trying to remember more difficult items before they were lost altogether. Subsequently, the remaining forty-nine subjects were asked for retrospective reports, and nine of these made a similar observation. This explanation is not quite so convincing in view of the fact that seven subjects offering it came from

the thirteen subjects remaining in the O-P group and only three from the twelve remaining in the G-N group, and it is this latter group that displays the paradoxical relationship most clearly. This final inconsistency in the data leads directly to an explanation of type 5.

The explanation to be offered rests on several premises:

(1) The probability of an item's being recalled at all depends on the degree to which it has attracted attention during presentation and been learned. In the present case three factors are of importance.

(a) Serial position. Material at either end of the list is more readily learned and recalled than material in the middle.

(b) Isolation (more accurately, structural change) increases the probability of recall.

(c) Item uniqueness. The ease with which a given item may be learned will depend on previous commerce with it, and/or some intrinsic property of the item itself.

(2) The order in which an item is recalled depends on:

(a) The degree to which it has been learned. Other things being equal, the better an item has been learned the earlier it will be emitted.

(b) The immediately preceding items during recall. Items of a given class tend to favour the early emission of other items in that class.

(3) Nouns 'belong together' for two reasons:

(a) They are recognizably of the same class.

(b) They have previously acquired associations.

Paralogs 'belong together' on only the former grounds.

Principles 1a and 2a have in fact been established by Deese (1957), Deese & Kaufman (1957), and Bousfield *et al.* (1958). For principle 1b *vide* Green (1958b). Principle 1c may be regarded as a formal statement of the self-evident. Principles 2b, 3a and 3b follow from what Bousfield *et al.* (1960) call 'associative clustering'.

We may now return to the original comparison between groups O-N and G-N. According to principles 1b and 2a an isolated item will be well learned and therefore tend to appear early during recall. This would produce the commonsense result of a simple correlation between probability of recall and order of recall.

In this particular case, however, there is an opposing tendency in that items from either end of the list will tend to be emitted early during recall (principles 1a and 2a), so that a paralog presented in the middle of a list of nouns will be displaced towards the end of the list. (Principles 2b, 3a, 3b.)

The commonsense result will therefore be attenuated by these special conditions, so that although the isolated item is well learned and has a high probability of being recalled, its order of recall will not match this high probability.

From the foregoing it follows that *office* will be emitted during recall along with the other nouns in an order corresponding to the degree to which each has been learned, while *gojey* will be shifted back towards the end of the recall list. That is, there will be a lack of correspondence between the high probability of *gojey's* being recalled and the order in which it is emitted.

This argument is supported by a further piece of qualitative data. Subjects in the G-N group frequently imported into their recall lists nouns that had not been presented. Sometimes it was clear that this had occurred on the basis of previously



learned associations. For instance, one subject, having correctly recalled *army*, followed this immediately with *escort*, an item that does not appear in the original list, but follows naturally enough in terms of sequential probabilities.

We would therefore appear to have arrived at a satisfactory explanation for Saltz & Newman's anomalous result. The differences between the O-N and O-P groups can also be readily accounted for. The differences in total recall and error scores follow from principle 3*b*. A comparison of the order of recall is complicated by the number of errors occurring in the O-P group. In the O-N group *office* is recalled about half way through, whether or not errors are taken into account. With the O-P group *office* is displaced slightly forward if errors are included, but much further back if errors are excluded. In other words, most of the invented items tend to be emitted after *office* while most of the genuine paralogs are emitted according to principle 3*a*.

A similar argument applies to a comparison of the O-P and G-P groups. *Office* is emitted earlier than *gojey*, which is to be expected from principles 1*c* and 3*b*, but this advantage is lost when errors are excluded from the analysis. *Gojey* then appears slightly earlier; 67% of the way through the correctly recalled items against *office* at 80%. This outcome is further complicated by the fact that although the mean number of items recalled (5.1 *v.* 5.4) is about the same if errors are included, there is a tendency to recall more correct items in the G-P group (1.8 *v.* 3.0). That is, the presence of *office* among a list of paralogs tends to interfere with the correct recall of the contextual items. This difference is significant at the 5% level (two-tail) using a Mann-Whitney *U* test. Just why paralogs should tend to be misspelt in the presence of a meaningful item is by no means clear. To a lesser extent *gojey* tended to upset the recall of correct nouns in the G-N group in an analogous manner.

The only question remaining to be considered is whether these findings affect Saltz & Newman's main conclusions. There seem to be grounds for asserting that they do. Their statement of the issue is as follows (1959, p. 446):

'The hypotheses to be tested in the present paper arise from the S-R association class of explanations for the von Restorff effect.

(1) According to this class of explanations, isolation of an item in a list increases the effective strength of the connexion between the isolated term and its intralist stimulus. Therefore, in the controlled association test the isolated item should be elicited by its appropriate intralist stimulus term more frequently than the control item is elicited by its appropriate stimulus.

(2) In terms of response generalization theory, the tendency for one response to intrude on another is assumed to be a function of the similarity between the responses. Since the non-isolated control item is similar to the other items of the list, while the isolated item is not, the isolated item should occur less frequently than the control item as an intrusion to inappropriate intralist stimuli.'

From their results, using a controlled association test, they conclude regarding their first hypothesis (1959, p. 447) that:

'These results do not support the hypothesis that isolation facilitates the association between the isolated item and its appropriate stimulus, after a single presentation of the serial list.'

This conclusion overlooks the fact that previously learned associations can account for an outcome such as theirs.

Regarding their second finding: their conclusion (1959, p. 450) that 'The primary effect of isolation appears to be an accelerated learning of the isolated item as a response' looks sound. The isolated item is certainly emitted frequently, and appears to be little stimulus-bound, regardless of whether it is meaningful or not. Deese & Marder (1957, p. 599) make a closely related point when they conclude that 'the context of items is forgotten much more readily than the items themselves.'

Although the evidence for their first hypothesis has not given the unequivocal support to their view that was intended, it should be pointed out that there are other quite independent grounds for regarding the serial interference type of explanation as inadequate. Work by Green (1958a, b) has shown that neither the Gestalt nor the serial interference account of the von Restorff effect can meet the facts.

Table 2. *Relationship between probability of recall and mean ordinal position of recall*

	Probability of recall	Ordinal position
Isolated	0.39	3.36
Massed	0.13	4.92
Comparison for items in fourth serial position.		
Isolated	0.24	3.95
Massed	0.10	5.89
Comparison for items in seventh serial position.		

It follows that when the appropriate controls are exercised, then the expected commonsense relationship between probability of recall and order of recall should hold. A quick glance back to an earlier experiment by Green (1956) tests this conclusion. In this experiment the presented lists were so designed that the critical items appeared in the same serial position whether massed or isolated, there were equal amounts of two kinds of material in each list, either kind being equally likely to have previously acquired associations, and the two kinds of material were so distributed throughout the list that neither would have an advantage in being recalled earlier than the other. Under these conditions items isolated in serial position 4 had a probability of being recalled of 0.39 and were emitted at a mean ordinal value of 3.36. When massed in the same serial position the same items had a probability of being recalled of 0.13 and were emitted at an ordinal value of 4.92 (see Table 2). In other words, the greater the probability of being recalled, the sooner the item is emitted. This result was confirmed by other groups in which the critical items were presented in the seventh serial position. The corresponding figures are also given in Table 2. In Saltz & Newman's experiment the isolated item got attention because it fell into a different *category* of meaningfulness from the surrounding material. Had they isolated *office* by presenting it, for example, in a different colour from the other nouns, instead of replacing it with a paralog, they would almost certainly have avoided the anomalous result examined here.

Thanks are due to Miss E. Hoyle for her skilful recruitment of subjects.

## REFERENCES

- BOUSFIELD, W. A. *et al.* (1956). The extension of Marbe's law to the recall of stimulus words. *Amer. J. Psychol.* **69**, 429-33.
- BOUSFIELD, W. A. *et al.* (1958). Serial position effects and the 'Marbe effect' in the free recall of meaningful words. *J. Gen. Psychol.* **59**, 255-62.
- BOUSFIELD, W. A. *et al.* (1960). Partial response identities in associative clustering. *J. Gen. Psychol.* **63**, 233-8.
- DEESE, J. (1957). Serial organisation in the recall of disconnected items. *Psychol. Rep.* **3**, 577-82.
- DEESE, J. & KAUFMAN, R. A. (1957). Serial effects in recall of unorganised and sequentially organised verbal material. *J. Exp. Psychol.* **54**, 180-7.
- DEESE, J. & MARDER, V. J. (1957). The pattern of errors in delayed recall of serial learning after interpolation. *Amer. J. Psychol.* **70**, 594-9.
- GREEN, R. T. (1956). Surprise as a factor in the von Restorff effect. *J. Exp. Psychol.* **52**, 340-4.
- GREEN, R. T. (1958*a*). Surprise, isolation, and structural change as factors affecting recall of a temporal series. *Br. J. Psychol.* **49**, 21-30.
- GREEN, R. T. (1958*b*). The attention-getting value of structural change. *Br. J. Psychol.* **49**, 311-14.
- HULL, C. L. *et al.* (1940). *Mathematico-Deductive Theory of Rote Learning*. Yale University Press.
- SALTZ, E. & NEWMAN, S. E. (1959). The von Restorff isolation effect: test of the intralist association assumption. *J. Exp. Psychol.* **58**, 445-51.

(Manuscript received 26 October 1961)



## TRANSFER BETWEEN DIFFICULT AND EASY TASKS\*

By D. H. HOLDING

*University of Leeds*

The existing literature on unequal transfer between difficult and easy tasks shows that it is impossible to predict the direction of optimum transfer on the basis of the *relative* difficulty of the two tasks. The first experiment explores an *absolute* difficulty hypothesis to the effect that transfer might be greater from a point of optimum difficulty, whether to an easier or to a more difficult task; changes in difficulty were achieved by controlled variation of the complication of pursuit tracking courses. However, after a week's practice on the twelve task conditions, it appeared that transfer was consistently most effective in the difficult-easy direction.

In a second experiment, subjects were transferred between tasks with easy and difficult course amplitudes, at both easy and difficult levels of complication. In this case, the easy-difficult order of practice was better with the simpler courses, while the complex courses favoured difficult-to-easy transfer. It is concluded that difficulty is not a useful category for the prediction of transfer efficiency, and that the solution lies in examining the skills involved. Explanations are outlined in terms of inclusion and error size constancy, although it is probable that many other factors play a part.

### I. INTRODUCTION

During the last decade, a large number of studies in transfer of training have been concerned with the variable of task difficulty. The interest in tasks of varying levels of difficulty lies in the fact that such transfer is usually asymmetrical; that is to say, learning a task which is more, or less, difficult than a subsequent transfer task is often more effective than transfer in the reverse order.

The question arises (i) whether either order of training, difficult-to-easy or easy-to-difficult, is consistently superior. Given an affirmative answer to this question, one may further ask (ii) which of the two orders of training is the better. Given a negative answer to the first question, one may attempt to predict (iii) on which occasions one or other is superior. The most common view has ignored the first question, assuming with regard to the second that difficult-to-easy transfer is basically superior. The intention of the present study is to show that this view is false, since question (i) received a negative answer; and to provide a basis for answering question (iii) by investigating possible asymmetrical transfer mechanisms.

In order to summarize existing data, a clear understanding is required of what is meant by asymmetrical transfer. Any instance of transfer is inferred from the relations between three obtained or implied learning curves: that of the first, or training task; that of the second, or transfer task; and that of the third, control task, in which the transfer task is carried out without the prior experience of the training task. In all of these learning curves, the initial score, the rate of learning and the final level of performance may all vary with task difficulty, giving rise to a profusion of possible measures.

If difficult and easy tasks offer different amounts of effectively possible learning, comparisons of the raw differences between transfer and control tasks are obviously unsuitable. Assuming, as do most transfer studies, that most interest attaches to the

\* This paper is based on a Ph.D. thesis for the University of Durham. The work was supported by the Flying Personnel Research Committee.

early stages of the transfer task, the most reasonable measure is that advocated by Gagné, Foster & Crowley (1948), in which the difference between the initial control score and the initial transfer is represented as a proportion of the difference between the initial and final control scores. A difference in these proportions for cases of difficult-easy and easy-difficult transfer may therefore be taken as evidence of asymmetry. Similarly, if the transfer data consist only of single scores which represent the learning of each task, an asymmetry in the proportions formed by dividing the control scores into the differences between transfer and control scores must be accepted. Owing to the variety of methods used in transfer experiments, a number of studies using other measures have also to be considered.

In order to avoid an extended section on the literature, while providing the facts necessary for attempting generalizations, a brief if ungainly solution lies in tabulation. Accordingly, Table 1 lists the available studies in terms of the direction of their results, together with an indication of the task, the variable used to define difficulty, and some qualifying remarks. Those studies headed 'inconclusive' are listed because, while relevant to the problem of asymmetrical transfer, the results have been negative.

It is immediately clear from inspection of the table that, in answer to our question (i), neither order of training is consistently superior. The easy-to-difficult results include all the target speed studies, the variation of control lag, the incidental finding of Poulton on reaction times and the atypical result of Kaestner and Grant on course complexity. In addition, those transfer studies using mainly perceptual variation have tended to produce easy-to-difficult outcomes; the difficult-to-easy training order which is required for direct comparison has not been included for obvious reasons. The difficult-easy section comprises the other forms of task complexity, and the response variables of gear ratio, handwheel size, control friction and pressure. The results on display-control variation are largely inconsistent, a possible explanation being that reversed relationships appear more natural in compensatory tracking.

It is not easy to force these findings into a rational mould, the classes appearing to be of mixed composition. Although Day's (1956) suggestions may be rephrased to the effect that perceptual variables form the nucleus of the easy-difficult studies, while the difficult-easy work centres on response variables, it is difficult to see why target speed should be a predominantly perceptual variable, and equally difficult to accept task complexity as a mainly response variable.

There are, in any case, two major obstacles in the way of an attempt to summarize the data in this fashion, among the many minor divergencies of experimental method and treatment. The first arises from wide differences in the amounts of learning involved, while the second is a function of what may be called the absolute difficulty of the experimental task. The difficulty over amounts of learning, which in practice are often minimal, arises from the fact that difficult and easy tasks may give rise to learning curves of different characteristics. Krueger (1947) has, for instance, shown positive acceleration for difficult tasks, with reverse inflexion for easier versions. This would often result in an easy-difficult effect for small amounts of practice, changing to difficult-easy superiority at greater mastery. In fact, little will be transferred if little has been learned, whether as a result of the amount of practice allotted or the amount of practice required by the nature of the task. Consequently, where the

<i>Easy-to-difficult superior</i>	<i>Authors</i>	<i>Task</i>	<i>Difficulty variable</i>	<i>Remarks</i>
	Ammons, Ammons & Morgan (1956)	Pursuit rotor	Target speed	Five minutes practice
	Kaestner & Grant (1956)	Pursuit tracking	Aperiodic <i>vs.</i> sine wave courses	Asymmetry only after shorter of two training periods
	Levine (1953)	Compensatory tracking	Control lag	Also effect of similarity
	Lincoln & Smith (1951)	Pursuit tracking	Target speed	Low speed better for medium and high; but medium training better for low
	Lordahl & Archer (1958)	Pursuit rotor	Target speed	Asymmetry only in early transfer
	North & Jeeves (1956)	Colour/position discrimination	Simultaneous <i>vs.</i> successive	Rats
	Poulton (1956)	Choice reaction	Number of choices	Not designed as transfer study
<i>Easy-difficult better than difficult-difficult</i>	Baker & Osgood (1954)	Pitch discrimination	Stimulus similarity	Gradual increase in difficulty best
	Belbin, Belbin & Hill (1957)	Weaving	Stimulus magnification	Other training methods
	Lawrence (1952)	Brightness discrimination	Stimulus similarity	Rats
<i>Difficult-easy superior</i>	Andreas, Green & Spragg (1954)	Tracking	Pursuit <i>vs.</i> compensatory; natural <i>vs.</i> unnatural	Asymmetry not conclusive
	Baker, Wylie & Gagné (1950)	Pursuit tracking	Gear ratio (rate of response)	Eight minutes learning time
	Cook (1937)	Pyramid puzzles	Number of disks	More practice on difficult problems
	Gerall & Green (1958)	Two-handed tracking	Control friction	Based on extrapolated control scores
	Gibbs (1951)	(i) Steering over studs	Course complexity	Learning incomplete
	Jones & Bilodeau (1952)	(ii) Compensatory tracking	Handwheel size	Result reversed for total learning time
	Szafran & Welford (1950)	Two-handed co-ordination	Course complexity	Second halves of practice compared
	Barch (1953)	Chain throwing	Use of bar and mirror	No learning curves
<i>Inconclusive</i>		Two-handed tracking	Target size and control reversal	Difficulty variables confounded
	Barch & Lewis (1954)	Two-handed tracking	Reversal of controls	Supports easy-difficult, but no direct comparison
	Briggs & Waters (1958)	Two-dimensional tracking	Degree of component interaction	No significant differences
	Browne (1954)	Controlling roll and pitch	Fixed <i>vs.</i> moving horizon	Recalculated percentages equal
	Gibbs (1954)	Compensatory tracking	Isometric <i>vs.</i> free control	No percentage difference
	Green (1955)	Two-handed tracking	Target size	Easy-difficult, but no difference in error scores
	Green <i>et al.</i> (1955)	Two-handed tracking	Reversal of controls	Time-on-target scores may hide difference
	Holland & Henson (1956)	Compensatory tracking	Quickened <i>vs.</i> unquickened systems	Results reversed at different training levels
	Maltzman, Smith & Brooks (1955)	Figure-of-eight tracking	Target speed	Slow-fast superior
	Morin (1951)	Pursuit rotor	Target size	Time-on-target scores
	Muckler & Matheny (1954)	Pursuit tracking	Pounds friction	No differences
	Nystrom, Morin & Grant (1956)	Three-light discrimination	Paced <i>vs.</i> unpaced	Difficult-easy, but scores not comparable
	Voss (1955)	Pedestal sight	Target brightness	Easy-easy better than difficult-easy



common task requirements are relatively simple, extended practice may be devoted to response refinement, and there may be greater transfer from the more 'difficult' version; while prior practice on the 'easy' version may be superior in the framework of more exacting task requirements, where the primary problem is to grasp what is to be done, rather than to execute the response.

This formulation embodies the perceptual-motor distinction suggested by the summary in Table 1, but derives its putative effect from the level of difficulty of the total task requirements, within which 'difficult' and 'easy' versions are located. In other words, although relative difficulty is an unacceptable predictor of the amount of transfer, differences in the absolute level of difficulty in conflicting experimental situations might be responsible for the discrepancies. If the 'easy' and the 'difficult' tasks of, say, the speed studies are both 'difficult' relative to the two conditions explored in course complexity, it is to be expected that the results will be inconsistent. Better 'easy-to-difficult' transfer in the first case, with better 'difficult-to-easy' transfer in the second, would imply the existence of a middle level of difficulty from which optimum transfer was obtained, regardless of the direction of relative difficulty. Above the optimum, too little may be learned for effective transfer, while too little may be available for learning below the optimum. Before it can be concluded that the concept of difficulty is otiose, or misleading, in the field of asymmetrical transfer, an absolute difficulty hypothesis of this sort must be considered.

In the absence of an accepted scale of difficulty, the best course appears to be to investigate the transfer relations obtaining over a range of difficulty wide enough to accommodate any likely optimum. This is attempted in Expt. 1, in a training situation sufficiently protracted to permit asymptotic learning.

## II. EXPERIMENT 1

### *Subjects*

The sixty subjects who completed the experiment were an 'as available' sample of R.A.F. personnel; their ranks varied from squadron leader to aircraftman, and their trades through air fitter, radar operator, mess steward, instrument mechanic and so on. Very few flying personnel could be spared for the length of the experiment.

### *Apparatus*

The equipment, described more fully elsewhere (Holding, 1961), provided for the pursuit tracking of a horizontally moving target spot by means of a small radar joystick. The joystick, mounted on an arm-rest, controlled a vertical cursor line at a gear ratio of 1:1, in the expected direction of motion relationship. The target spot, and the cursor line, were displayed on a double beam oscilloscope, and tracking error was electronically squared, integrated and displayed on a dekatron counter.

The target courses were random in pattern (see Fig. 2) and were derived from the output of a 'white noise' generator, severely curtailed by low-pass filters. White noise may be regarded as the sum of an infinite series of sine waves, equal in amplitude but random in phase, resulting in a random wave form whose instantaneous amplitudes are normally distributed. For human tracking purposes, all frequencies higher than a few cycles per second are eliminated, so that any given target course will thus contain excursions at only the frequencies within its pre-set bandwidth; both the bandwidth, or range of included frequencies, and the average amplitude, were variable.

With a change in amplitude, the average excursion of the target becomes larger or smaller, while the degree of complication remains constant; with a change of bandwidth, the course varies in complication, although the average amplitude remains the same.

Testing was carried out in a darkened room, and, with seating adjusted, subjects viewed the oscilloscope through a visor which kept viewing distance constant at about 40 cm.

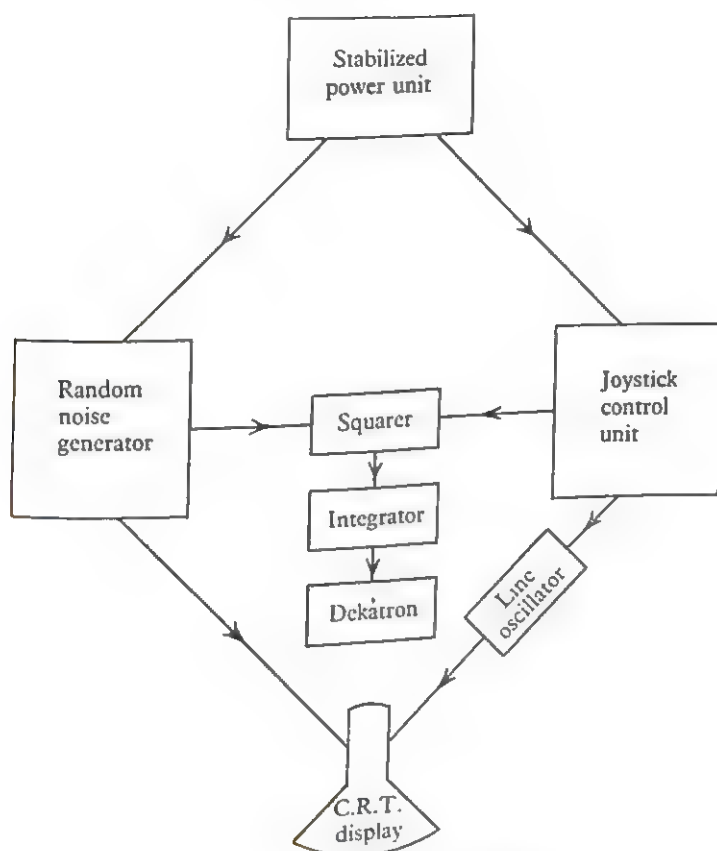


Fig. 1. Block diagram of the apparatus.

Bandwidth: 1 cyc./sec.

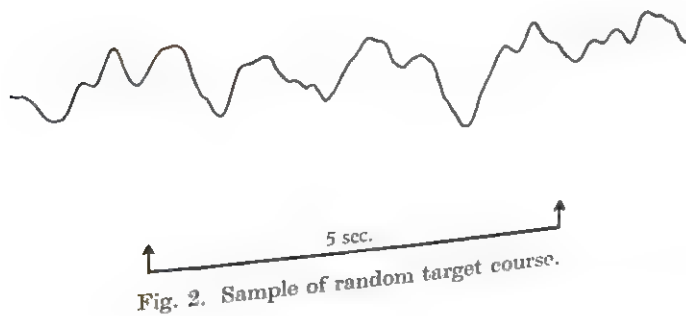


Fig. 2. Sample of random target course.

#### Design

Four courses were chosen on the basis of previous work, at a common root mean square (R.M.S.) amplitude of 10 mm. at the oscilloscope face. These ranged from (1) the clearly easy bandwidth of 0-0.29 cycles per second (cyc./sec.), to (4) the extremely difficult bandwidth of 0-1.16 cyc./sec., at

intervals of 0.29 cyc./sec. Courses easier than the lowest bandwidth, where the R.M.S. error expected of practised subjects was well under half the course displacement, would present scoring difficulties; while, since the practised error for course (4) is approximately 90% of the course displacement, there is no doubt concerning its difficulty. Subjects were transferred in all possible orders between the four courses, there being thus twelve conditions of transfer:

1-2	2-1	3-1	4-1
1-3	2-3	3-2	4-2
1-4	2-4	3-4	4-3

It will be seen that the results for all three groups of subjects practising on a given first course may be pooled to form control learning curves, against which the transfer results may be separately compared.

### Procedure

Subjects spent five days on the first course, and four days on the second, each subject tracking the appropriate course for half an hour each day. Tracking sessions consisted of six 2 min. runs interspersed with 2 min. rest pauses, so that each subject received thirty runs on the first task, and twenty-four transfer runs. By testing twelve subjects each day, the experimental programme was resolved into fortnightly groups of twelve subjects, each subject in the group being allotted to a different experimental condition.

At the outset, the action of the scoring apparatus was demonstrated, and subjects were told that the task was to follow the moving spot as closely as possible. Each subject was given a minute's preview before starting to track his first or second course, and knowledge of score was given after each run.

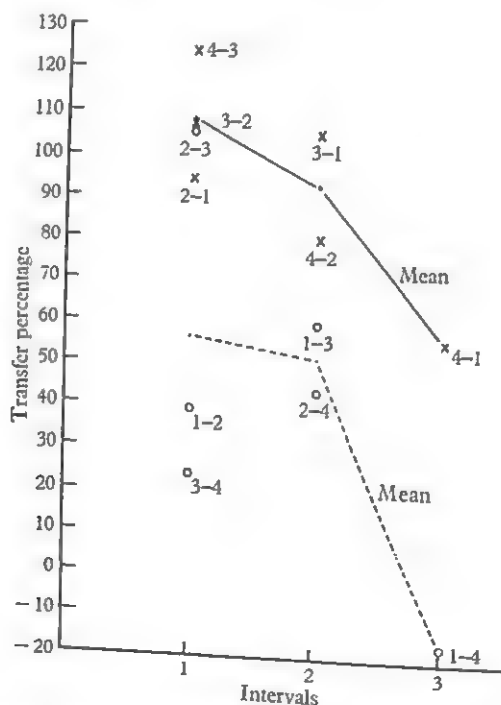


Fig. 3. Transfer as a function of the direction of change of difficulty, and of the interval between target courses. Each target course is designated by a number (the higher the number, the more difficult the course). x—x, Difficult-easy; o---o, easy-difficult.

### Results

Learning curves plotted from the square roots of the 2 min. dekatron scores showed that the level of performance had apparently stabilized by the end of the first week; courses 2, 3 and 4 gave approximately equal amounts of learning, the improvement on



course 1 being about half as much. Transfer percentages were obtained from the group means, using error in the first half session as the initial and transfer scores, and all six runs of the last day's scores for the terminal control scores. Contrary to hypothesis, a clear difficult-easy effect was obtained: comparison of the six transfer percentages in the difficult-easy direction, with the six percentages in the reverse direction, shows that the difficult-easy order is in every case superior. At the same time, the traditional effect of similarity also appeared, in that grouping the percentages by intervals between courses shows decreasing transfer with increasing differences in difficulty. Both these effects are seen in Fig. 3, where the percentages are plotted against the number of courses by which the transferred task differed from the control task, with direction of relative difficulty as the parameter.

To provide data for analysis, the individual transfer percentages were computed for each subject. An analysis of variance, summarized in Table 2, shows the direction effect to be significant. The differences between intervals are only 'significant' at the 10% level; but bearing in mind the prior evidence on similarity, it is probable that this represents a genuine effect.

Table 2. *Analysis of variance: individual transfer percentages*

Source	Sum of squares	D.F.	Mean square	F	Significance
Between directions	40,093.35	1	40,093.35	8.13	$P < 0.01$
Between intervals	29,480.68	2	14,740.34	2.99	N.S.
Directions $\times$ intervals	1,482.08	2	741.04	< 1	N.S.
Residual	274,800.07	54	5,088.89	—	—
Total	345,856.18	59	—	—	—

The pattern of transfer relationships was essentially the same at the conclusion of the second week, although differences in performance between the experimental groups were naturally reduced.

#### Discussion

Since there appears to be no evidence to support the absolute difficulty hypothesis, while the effects of relative difficulty have been shown to be inconsistent, it is clear that the concept of task difficulty can contribute little to the understanding of asymmetrical transfer. An explanation must presumably be sought in the detailed analysis of the skills investigated, and difficulty may be rejected as an accidental attribute of those task characteristics which do play a significant part in determining transfer.

In the case of the foregoing results, which align themselves with the previous findings on complexity, a ready explanation lies in the principle of inclusion. A high bandwidth course includes low frequency components, which have therefore been practised during the training on the 'difficult' task. In transferring from low to high bandwidths, however, the subject is confronted with higher frequency components which have not previously been met, and the result therefore favours the difficult-easy direction. The change in timing necessitated by the bandwidth changes tend to preclude attention to precision of response.

The other variable in the present task is that of course amplitude, and within the context of the constant timing of a single bandwidth, another mechanism may be demonstrated. Where the speed or timing requirements permit attention to accuracy of response, it may be suggested that there is some tendency to carry over from the

training task, not the mean level of response, but the mean error size. An error which is small relative to a wide amplitude course, carried over intact to the new course, will become a relatively large error in tracking a lesser amplitude; and conversely, an error at low amplitude becomes relatively smaller when transferred intact to a higher amplitude course. The result of such a transaction is to raise the transfer value of training on the easy task, compared with training on the more difficult version. Such a result was in fact obtained in an earlier pilot study in which course amplitude was varied. However, the error size explanation depends on the capacity of the subjects to track accurately to the standards evolved during the training period. While both courses are at a low level of complexity, this condition will obtain; but at higher bandwidths, where tracking is in any case more difficult, subjects will be unable to achieve close tolerance objectives, and this transfer mechanism will break down. It may therefore be expected that the degree of asymmetry will be considerably reduced. This hypothesis is tested in Expt. 2, where transfer between amplitudes is investigated at two bandwidth levels.

### III. EXPERIMENT 2

#### *Subjects*

The twenty-four subjects were volunteers from the first-year psychology course in Newcastle. Ten of these were male, and fourteen female.

#### *Design*

The subjects were assigned to four groups, transfer taking place in both directions between the two R.M.S. amplitudes (7.3 and 10 mm.), at each of the two bandwidths (0.044 and -0.87 cyc./sec.). The first-task scores for each group were again the control scores for the group transferring to that task.

#### *Procedure*

Experimental sessions lasted half an hour per day for each of five days. During each half-hour session, six 2 min. runs, interspersed with rest pauses, were carried out. In order to shorten the subjects' task, it was decided to omit the data on the later course of second-task learning, so that each subject did twenty-seven 2 min. runs on the appropriate training course, followed by three runs on the other amplitude at the same frequency bandwidth. Other points of procedure, and the apparatus, were as previously described.

#### *Results*

The transfer percentages, based on the means for the first and last three control runs, and the three transfer runs, are listed in Table 3.

Table 3. *Percentage transfer between amplitudes in Expt. 2*

	Easy-difficult	Difficult-easy	Mann-Whitney U	Significance
Low bandwidth	141	73	1	$P < 0.01$
High bandwidth	65	121	5	$P < 0.05$

It will be seen that 'easy-difficult' transfer appears superior at the lower bandwidth; while, unexpectedly, the reverse is true at the higher bandwidth. This conclusion is verified by a two-tailed Mann Whitney test on the individual transfer percentages, the differences being significant at each bandwidth. The expected decrease in asymmetry was therefore exceeded, in a clear instance of both directions of asymmetry in a single experimental framework.

## IV. DISCUSSION

Where the more difficult task differs from its easier version in complexity, so that, as with the bandwidth variable of Expt. 1, the more difficult version in some sense contains more than the easier, the principle of inclusion clearly favours the difficult-to-easy order. If inclusion may be implicit, rather than literal, such a principle may be extended to cover the response variables reviewed above. Since the including and the included tasks may not contain physically 'identical elements', it may have to be recognized that skills or skill components exist, the learning of which renders unnecessary the learning of less demanding tasks. This may apply to whole skills like the chain-throwing of Szafran & Welford (1950), or to discrete components such as responses to differing handwheel size or control friction.

On the other hand, the effect of any tendency to error size constancy, which has been taken to favour the easy-difficult order, will to some extent depend upon the inclusion or exclusion of high-error components. Thus, if error adversely affects the higher frequency components which do not occur in the narrower bandwidths, subjects transferring in the high-low direction will be equipped with practice at, and reasonably small error expectancies for, the lower frequencies, so that any easy-difficult advantage will be reduced.

The interaction between inclusion and error size constancy will be further complicated by the strategy adopted by the subject, and by the stress upon speed or accuracy demanded by the structure of the task. In a tracking task, a subject may opt to anticipate, or to follow as closely as possible; and while anticipation may be optimal with an easy course, only following is practicable with a complex course, so that the transfer of these strategies is another source of asymmetry. Similarly, since subjects may generally sacrifice accuracy for speed, and will be called upon to do so in varying amounts by different tasks, it is likely that the set engendered by difficult and easy versions will again produce mutual transfer relationships which tend to asymmetry. The network of interacting factors is thus extremely complex, and solutions can only be found for narrow classes of task.

It is to be expected that the more perceptual tasks will emphasize anticipation and accuracy, rather than speed of following, and an easy-difficult result will be predicted in general. Very often, however, this conclusion will depend less on the amount of error committed in a difficult version, than on the possibility of any learning taking place. The other main group of easy-difficult results, the speed studies, may certainly bear an error size explanation. Variation of speed took place in the context of a basically uncomplicated task, which remained the same for all the versions—unlike the present bandwidth results, in which new frequencies were successively introduced. Anticipation was relatively simple, so that, paradoxically, speed of readjustment did not differentiate between courses, and the inclusion relationship was effectively one of equivalence. In these circumstances, the standards of accuracy evolved during training may be considered crucial. It should be noted that this account does not preclude the possibility that, with precision held constant, a practised movement on a faster course may more easily be slowed than a slower one may be hastened, over the range of speeds for which a basic movement pattern is maintained.



It is clear, therefore, that the use of the concept of difficulty must give way to far more detailed analysis of the appropriate skill, if asymmetrical transfer is to be successfully predicted.

It may be true, for instance, that skill at longer control lags in some sense includes skill at shorter lags, but the errors committed at the shorter lags will tend to be smaller, so that both inclusion and error size mechanisms will be involved. Further, the difference in error size may be smaller after prolonged practice, while the acquired timing mastery may permit greater attention to anticipation with the shorter lags; so that the amount of practice, and the speed-accuracy or other set implied by the task structure and subject expectancies will interact with the other factors. We may therefore expect that transfer between lags under various conditions will give conflicting transfer results. What we may not say, it is evident, is that these effects occur merely because a task with lag is more difficult than a task without such lag.

### REFERENCES

- AMMONS, R. B., AMMONS, C. H. & MORGAN, R. L. (1956). Transfer of skill and decremental factors along the speed dimension in rotary pursuit. *Percept. & Mot. Skills*, **6**, 43.
- ANDREAS, B. G., GREEN, R. F. & SPRAGG, S. D. S. (1954). Transfer effects between performance on a following (modified SAM two-hand co-ordination test) and a compensatory tracking task (modified two-hand pursuit test). *J. Psychol.* **37**, 173-83.
- BAKER, K. E., WYLLIE, R. C. & GAGNÉ, R. M. (1950). Transfer of training to a motor skill as a function of variation in rate of response. *J. Exp. Psychol.* **40**, 721-32.
- BAKER, R. A. & OSGOOD, S. W. (1954). Discrimination transfer along a pitch continuum. *J. Exp. Psychol.* **48**, 241-6.
- BARCH, A. M. (1953). The effect of difficulty of task on proactive facilitation and interference. *J. Exp. Psychol.* **46**, 37-42.
- BARCH, A. M. & LEWIS, D. (1954). The effect of task difficulty and amount of practice on proactive transfer. *J. Exp. Psychol.* **48**, 134-42.
- BELBIN, E., BELBIN, R. M. & HILL, F. (1957). A comparison between the results of three different methods of operator training. *Ergonomics*, **1**, 39-50.
- BRIGGS, G. E. & WATERS, L. E. (1958). Training and transfer as a function of component interaction. *J. Exp. Psychol.* **56**, 492-500.
- BROWNE, R. C. (1954). Figure and ground in a two-dimensional display. *J. Appl. Psychol.* **38**, 462-7.
- COOKE, T. W. (1937). Amount of material and difficulty of problem-solving. *J. Exp. Psychol.* **20**, 288-96.
- DAY, R. H. (1956). Relative task difficulty and transfer of training in skilled performance. *Psychol. Bull.* **53**, 160-8.
- GAGNÉ, R. M., FOSTER, H. & CROWLEY, M. E. (1948). The measurement of transfer of training. *Psychol. Bull.* **45**, 97-130.
- GERALL, A. A. & GREEN, R. F. (1958). Effect of torque changes upon a two-handed co-ordination task. *Percept. & Mot. Skills*, **8**, 287-90.
- GIBBS, C. B. (1951). Transfer of training and skill assumptions in tracking. *Quart. J. Exp. Psychol.* **3**, 99-110.
- GIBBS, C. B. (1954). The continuous regulation of skilled response by kinaesthetic feedback. *Brit. J. Psychol.* **45**, 24-39.
- GREEN, R. F. (1955). Transfer of skill on a following tracking task as a function of task difficulty. *J. Psychol.* **39**, 355-70.
- GREEN, R. F., GOODENOUGH, D. R., ANDREAS, B. G., GERALL, A. A. & SPRAGG, S. D. S. (1955). The effects of phases of rotation of control cranks on performance levels and transfer effects in compensatory and following tracking. *U.S. Speedeven.* No. 241-6-22.
- HOLDING, D. H. (1961). *Transfer of learning between motor tasks of different levels of difficulty*. Unpublished Ph.D. thesis, University of Durham.

- HOLLAND, J. G. & HENSON, J. B. (1956). Transfer of training between quickened and unquickened tracking systems. Washington, *N.R.L. Rep.* No. 4703.
- JONES, E. I. & BILODEAU, E. A. (1952). Differential transfer of training between motor tasks of different difficulty. *H.R.R.L. Bull.* No. 52-35.
- KAESTNER, N. F. & GRANT, D. A. (1956). Transfer of training in tracking as a function of the predictability of uni-dimensional target courses. *J. Gen. Psychol.* **55**, 103-16.
- KRUEGER, W. C. F. (1947). Influence of difficulty of a perceptual-motor task upon acceleration of learning curves. *J. Educ. Psychol.* **38**, 51-3.
- LAWRENCE, D. H. (1952). The transfer of a discrimination along a continuum. *J. Comp. Physiol. Psychol.* **45**, 511-16.
- LEVINE, M. (1953). Transfer of tracking performance as a function of a delay between the control and the display. *W.A.D.C. Tech. Rep.* No. 52-237.
- LINCOLN, R. S. & SMITH, K. U. (1951). Transfer of training in tracking performance at different target speeds. *J. Appl. Psychol.* **35**, 358-62.
- LORDAHL, D. S. & ARCHER, E. J. (1958). Transfer effects on a rotary pursuit task as a function of first task difficulty. *J. Exp. Psychol.* **56**, 421-6.
- MALTZMAN, I., SMITH, W. M. & BROOKS, L. O. (1955). Some effects of different training conditions and manifest anxiety upon target tracking. *Percept. & Mot. Skills*, **4**, 185-91.
- MORIN, R. E. (1951). Transfer of training between motor tasks varying in precision of movement required to score. *Amer. Psychologist*, **6**, 390 (Abstract).
- MUCKLER, F. A. & MATHENY, W. G. (1954). Transfer of training in tracking as a function of control friction. *J. Appl. Psychol.* **38**, 364-7.
- NAMIKAS, G. & ARCHER, E. J. (1960). Motor skill transfer as a function of intertask interval and pretransfer task difficulty. *J. Exp. Psychol.* **59**, 109-12.
- NORTH, A. J. & JEEVES, M. (1956). Inter-relationships of successive and simultaneous discrimination. *J. Exp. Psychol.* **51**, 54-9.
- NYSTROM, C. O., MORIN, R. E. & GRANT, D. A. (1956). Transfer effects between automatic and unpaced schedules in a perceptual-motor task. *J. Gen. Psychol.* **55**, 9-17.
- POULTON, E. C. (1956). The precision of choice reactions. *J. Exp. Psychol.* **51**, 98-102.
- SZAFRAN, J. & WELFORD, A. T. (1950). On the relation between transfer and difficulty of initial task. *Quart. J. Exp. Psychol.* **2**, 88-94.
- VOSS, J. F. (1955). Effect of target brightness and target speed upon tracking efficiency. *J. Exp. Psychol.* **49**, 237-43.

(Manuscript received 29 September 1961)





## PERIPHERAL VISION, REFRACTORINESS AND EYE MOVEMENTS IN FAST ORAL READING

By E. C. POULTON

*Medical Research Council Applied Psychology Research Unit, Cambridge*

Naval ratings read typescript aloud through a window whose speed and size varied systematically in different trials. Eye movements were recorded electrically. Errors increased significantly as the window was reduced from a full line to five words, probably as a result of the tighter pacing and reduced peripheral vision. Regressions were unaffected. This is a serious criticism of reading films using windows to 'improve' the reading of adults.

Errors rose steeply to 40% as the time for which letters were exposed fell from 0.3 to 0.2 sec.; fixations did not increase in frequency. Thus in reading the visual system behaves like a single-shot camera firing at not more than about five times per second.

### INTRODUCTION

#### (i) *A moving window as a training device*

Films designed to increase the speed of reading often project in focus or with adequate brightness only a limited number of words at a time (Poulton, 1961). In the Harvard films (Perry & Whitlock, 1948) the window jumps along the line every 0.30 or 0.45 sec. In the Carborundum films (1959) the window moves smoothly along the line except for a stationary period at the start and end. Curiously enough, no research appears to have been carried out to discover to what extent constraints of this kind interfere with reading, nor has the best size of window been determined objectively. In selecting their smallest window Perry & Whitlock simply presented various sizes to their poor readers, and found the smallest with which they reported that they were at all comfortable (1954, p. 25). The Carborundum Company appear to have adopted Perry & Whitlock's standards without question.

Exps. I and IV show the effects of reducing the size of the window. Exp. I also shows that prose, despite its larger syntactical and meaningful units, is no more upset by a relatively small reduction than are scrambled prose and comparably spaced random letters.

#### (ii) *Single-shot and ciné camera models*

The single-shot and the ciné camera as alternative models of the visual system help in understanding the possible effects of windows smaller than those used in training films. The single-shot camera has a rate of firing of not more than about five times per second, corresponding to the maximum frequency of fixations in reading (Woodworth, 1938, pp. 722-5). The upper limit also corresponds to the duration of about 0.2 sec. after the computing mechanism in the brain has accepted information before it can be cleared for the next intake (Poulton, 1950; Welford, 1959). The ciné camera fires more rapidly; an upper limit might be set at about 20 times per second, since with normal lighting the apparent brightness of a display is reduced with exposures shorter than about 0.05 sec. (Woodworth, 1938, p. 565). The camera is connected to the central computer through a sensory store which holds the information until the computer can accept it.

Both models predict that errors in reading will increase markedly as the time for which each letter is exposed is reduced to the minimum average time interval between shots of the camera. They differ in the duration of the exposure at which this will occur. Fig. 1 illustrates the display every 0.05 sec. through a window of four letterspaces moving at 20 letterspaces per second, or 240 words per minute. If the single-shot camera fired regularly every 0.2 sec., it might take the views printed in bold letters. Every letterspace would then be registered once, but without any overlap between successive views. In practice there is bound to be some variability in the periodicity of the camera; thus it might fire at 0.0, 0.15 and 0.40 sec. If it did so, the letter *s* of 'first' would be registered in two shots, while the letter *a* of 'was' would fail to be registered at all. This would be likely to interfere considerably with fast reading, and would occur as soon as each letter was exposed for an average of only about 0.2 sec.

In contrast, if the ciné camera fired regularly every 0.05 sec., it would take all the views in Fig. 1, and there would be no chance of a letter failing to be registered. This model predicts that reading will not start to break down until each letter is exposed for considerably less than 0.2 sec.

Time (sec.)	Display
0.00	<b>f</b> i r s
0.05	<b>i</b> r s t
0.10	<b>r</b> s t
0.15	<b>s</b> t w
0.20	<b>t</b> w a
0.25	w a s
0.30	w a s
0.35	a s t
0.40	<b>s</b> t h

Fig. 1. Successive views through a moving window of four letterspaces taken by a hypothetical ciné camera firing every 0.05 sec. The bold letters show the views taken by a single-shot camera firing only every 0.2 sec.

### (iii) *Silent and oral reading*

The experiments used oral rather than silent reading to enable the experimenter to tell precisely where the reader ran into difficulty. This is not possible with silent reading; the experimenter has to rely upon a delayed and probably inaccurate report from the reader, supplemented only by the results of a rather imprecise test of comprehension. However, in the present experiments oral reading resembled silent reading. The subject was told to gabble the words as quickly as he could, not to pay attention to punctuation, and not to correct misreadings. Rates of reading of up to at least 300 words per minute were thus investigated, at which rate the eyes had to move along the line at up to 400 words per minute owing to the way in which the apparatus worked. The fastest reader spoke at 425 words per minute. These rates are well within the range of the rates of students classified as good silent readers by Anderson & Swanson (1937). Thus although the results may be typical only of fast oral reading, it seems probable that they apply also to silent reading other than the kind in which attention is not paid to every single word, and which therefore merges into skimming.

## EXPERIMENT I. LARGE WINDOWS

Exp. I was designed to show the effect of reducing the size of the window while holding its speed constant for any one individual. It answers the questions: (a) How far can the window be reduced before reading is appreciably upset? (b) Is the reading of prose upset by larger windows than the reading of scrambled prose or of comparably spaced random letters?

*Method*(i) *Experimental subjects*

A different twelve subjects performed each part. All were naval ratings aged between 18 and 31 years. Most had left school at 14 or 15 and had been apprentices or had taken miscellaneous jobs before joining the Royal Navy. One rating, who could only read with difficulty, had to be rejected.

(ii) *Apparatus*

The carriage of a typewriter held a page of double-spaced typing. A clear celluloid disk rotated 0.4 in. above the carriage, and had attached to it a mask of black paper containing a spiral window. The disk and mask were covered by a screen, which contained a left-to-right slit of sufficient dimensions to show a single type written line. The whole apparatus was tilted forward 20° like a lectern. The page was illuminated from the back by daylight and a 60 W. lamp. At the lowest level of illumination which occurred in the late afternoon, the page reflected 5 to 8 foot lamberts in an otherwise darkened room, which is quite adequate for reading (Tinker, 1943).

When the page was viewed through the slit, the spiral window exposed a section of typewritten line of a fixed size which moved smoothly across the page from left to right. Once it had passed beyond the right-hand edge of the page there was a short delay before it came on again across the left-hand edge. During the delay the carriage of the typewriter was moved on smartly by a solenoid to the next typewritten line. In addition to this delay, during which no letters were visible, when a window first appeared on the left its leading edge had to move some distance along the line before it attained its full size. Similarly when the window reached the end of the line it gradually became smaller before it finally disappeared. With prose the narrowest window of three letterspaces was filled with letters for 72% of the time. For this same percentage of time the widest window of 79 letterspaces exposed at least 12 letters, a window 37 letterspaces wide exposed at least 8 letters, and a window 9 letterspaces wide exposed at least 4 letters. With all sizes of window the rate at which the eyes had to move across the middle section of the line was thus about one-third quicker than the rate at which the material had to be spoken.

A microphone suspended close to the subject's face picked up what he said, and fed it into a Ferrograph magnetic-tape recorder.

(iii) *Reading material*

The material to be read was typed in élite lettering, 12 letterspaces to the inch. The lines of prose ranged in length from 63 to 70 letterspaces. Only one space separated adjoining words, a punctuation mark when present occupying this space. The prose was taken from the descriptive parts of H. G. Wells's *The Time Machine*, in Basic English (1951). Practice passages averaged about eighty lines, and test passages fifteen lines. The average length of words was four letters and the mean reading-ease score was 76 (Flesch, 1948). The passages of *scrambled prose* were identical with the passages of connected prose, except that within lines the order of words was randomized, and punctuation marks and capital letters were omitted. *Random letters* were typed in lines of ten, each letter separated from the next by six blank spaces. Each practice and test passage was always preceded by five lines of alternate numerals and dashes. The first line contained 5s, the second line 4s—and the last line 1s. These lines enabled the subject to become adjusted to the size and rate of movement of the window before he had to start reading.



(iv) *Design*

The experimental conditions are listed in Tables 1 and 2.

*Exp. IA* used prose and random letters with six sizes of window. Six subjects read the prose on the first day and the random letters on the next day; the other six subjects took the conditions in the reverse order. For each subgroup the six sizes of window, six test passages, and the order of presentation were arranged in a Graeco-Latin square, which was the same for both types of material. The second subgroup was presented with the sizes of window in the reverse order to the first subgroup.

*Exp. IB* used scrambled prose as well as connected prose and random letters, and only four sizes of window. The different materials were presented on different days according to a balanced Latin-square design (Edwards, 1951). The four experimental conditions with each type of material were arranged in the same Graeco-Latin square.

(v) *Procedure*

The subject was instructed to read aloud as quickly as possible. He was told to breathe and move his eyes back from right to left while the typescript was invisible. With narrow windows he was warned to fixate the typed page, not the edge of the window. (Fixating the edge of the window gave the subjective impression of the window remaining stationary while the typescript moved in the opposite direction. All the letters were then clearly visible, but it was not possible to read.) The subject was also instructed not to attempt to correct misreadings, nor to worry about punctuation marks if present, lest he fall behind. He was told that samples of his reading would be played back to him after practice trials.

Before each part of *Expt. IA* there was a practice trial with each size of window. After this the subject read with the window of 79 letterspaces while its rate was adjusted between trials until he was making as nearly as possible 10% of errors. Each practice and experiment together took about 90 min. Before *Expt. IB* each practice was similar, except for a criterion of 5% of errors.

(vi) *Scoring and calculations*

While the record was played back, the experimenter noted three types of errors on a duplicated sheet: (a) omissions; (b) misreadings, when a word or letter was pronounced so that it could not be recognized, or a different word or letter was substituted; and (c) insertions, when an additional word, letter, or sound was inserted. A mistake which was followed directly by the correct word or letter counted as an insertion.

Statistical significance was generally assessed by *t*-tests. In the Runs test (Siegel, 1956, p. 52) a value of *z* was computed for each subject with each kind of experimental material. The *z*s for each material were summed over subjects, divided by  $\sqrt{N}$ , and the value of the resultant checked for probability against the normal distribution. Two-tailed tests were always used.

*Results*

Table 1 shows the mean rates of reading. The standard errors (s.e.) indicate the sizes of the individual differences. In *Exp. IB*, where the criterion was about 5% of errors with the largest window, the words of scrambled prose were read on average significantly more slowly than random letters, and only two-thirds as quickly as connected prose. Considering the differences in material and experimental method, the average rates correspond reasonably well to those reported by Pierce & Karlin (1957, Table 5, and p. 509).

Table 2 shows the proportion of errors with different sizes of window. The first reliable increase in errors with prose occurred in *Exp. IB* between windows of 13.2 and 5.0 words. A One-Sample Runs test (Siegel, 1956, p. 52) revealed a significant tendency for errors to occur in bunches. With the window of 79 letterspaces the

pooled data from the two experiments gave  $P < 0.01$  or better with each of the three kinds of material.

Out of the total errors with the moving window of 79 letterspaces, there was a significantly greater proportion of misreadings with the scrambled prose than with the connected prose ( $P < 0.02$ ): 60% of the misreadings of the scrambled prose made

Table 1. Mean rates of reading in Exp. I

Expt.	Prose (words per minute)		Scrambled prose (words per minute)		Random letters (letters per minute)	
	Mean	S.E.	Mean	S.E.	Mean	S.E.
1A	293*	15.5	—	—	177*	6.0
1B	219*	14.2	149†	7.9	162*†	6.0

\* Prose—random letters  $P < 0.01$  or better.

† Scrambled prose—random letters  $P < 0.02$ .

Table 2. Errors per 100 words or random letters in Exp. I

Size of moving window (letter- spaces)	Table 2. Errors per word													
	Prose						Scrambled prose		Random letters					
	No. of words exposed	Expt. 1A		Expt. 1B		Expt. 1B	Expt. 1B	No. exposed	Expt. 1A		Expt. 1B			
		Mean*	S.E.	Mean†	S.E.				Mean‡	S.E.	Mean§	S.E.	Mean§	S.E.
79								10	7.9	1.9	4.8	0.8		
37	13.2	12.8	1.6	5.5	0.7	6.3	0.8	5.3	10.8	1.7	—	—		
25	7.4	13.1	1.6	—	—	—	—	3.6	11.8	1.5	6.5	1.1		
13	5.0	14.1	1.6	8.7	1.1	8.2	1.1	1.9	16.0	2.1	12.1	1.4		
9	2.6	16.0	1.5	9.0	1.2	8.9	1.0	1.3	20.2	1.8	12.7	2.5		
5	1.8	20.8	2.3	13.9	1.9	11.8	1.9	0.7	28.0	2.3	—	—		
	1.0	31.1	1.8	—	—	—	—							

\* 2 words at the 0.01 level.

\* 13.2 words not different from 2.6 words, but different from 1.8 words at the 0.01 level.

† 13.2 words different from 5.0 and 2.6 words at the 0.05 level, and from 1.8 words at the 0.01 level.

‡ 13.2 words not different from 5.0 words, but different from 2.6 words at the 0.05 level and from 1.8 words at the

0.02 level.

§ 10 random letters not different from 3.6 letters, but different from 1.9 letters at the 0.01 level or better.

Not different from the comparable mean in Table 3.

better sense in their contexts than the correct words. It made no difference to the error rate whether the connected prose was read before the scrambled prose or afterwards. With all three types of material omissions tended to increase faster than misreadings as the total proportion of errors increased. Insertions always remained at 1–2% of the words or letters presented.

### Discussion

The following discussion is confined to the effect of windows comparable in size with those used in training films. Table 2 shows that all three kinds of material gave rather similar increases in errors as the size of the window was reduced. The correspondence between scrambled prose and random letters was to be expected, since a word of scrambled prose is almost an independent unit like a random letter, and the random letters were spaced out along the line like the words. But the prose was expected to suffer, with relatively large windows having little effect upon scrambled prose. The

failure to find this suggests that the increase in errors with the prose was determined very largely by factors not related either to its meaning or to its syntactical structure.

One such factor was the tighter pacing of the sensory processes produced by reducing the size of the window. When the subject stumbled in his reading, the smaller the window the more likely it would be that his view would be cut off before he had time to recover fully. This would not have affected his error rate when he had said the wrong word, since misreadings were counted as errors even when corrected. But when he was simply held up the tighter pacing might well have been a handicap.

Another possible factor was the reduction in peripheral vision. Owing to the 'sharp drop in visual activity in the parafoveal regions' (Luckiesh & Moss, 1942, p. 144) the recognition of words and letters is limited to the fovea and the retina closely surrounding it. The use of this area would not be restricted at a reading distance of 18 in. until the window was reduced to about 13 letterspaces, since a window of this size subtends an angle of about  $3^\circ$  at the eye, as compared with the  $1^\circ 10'$  of the fovea. However, it is possible to make out in peripheral vision the shape of approaching words, or of approaching letters in the case of spaced random letters; the distance ahead over which this could be done would be limited by the size of the window.

The Harvard (Perry & Whitlock, 1948) and Carborundum (1959) reading films start by presenting prose with three words of average length visible at a time through their moving window. The present evidence suggests that this is too little. In Exp. I B a significant increase in errors occurred with the prose when the window was reduced only as far as five words. It seems probable that some slight deterioration will occur as soon as any appreciable restriction is placed upon vision. This is the penalty which has to be paid for reading with a device which is designed to exert control over the direction of fixations.

## EXPERIMENT II. SMALL WINDOWS

Exp. II was designed to show the interaction between the size of the window and its speed of movement. It answers the questions: (a) Can the single factor of time for which each letter is exposed account satisfactorily for the effects of the changes in both independent variables? (b) For how long must each letter be exposed in order to prevent reading from breaking down?

### Method

The twelve new *experimental subjects* were all naval ratings aged between 17 and 22 years. Only connected prose was read.

The *design* is indicated in Table 3. The four sizes of window each moved at three out of five rates. A Latin square design was used for these twelve conditions, twelve test passages and twelve subjects. Each subject performed all conditions with one size of window before proceeding to the next size. The windows were always presented in order of size, and the speeds in order of speed. Six subjects started with the largest window, and six with the smallest. Of each six, half were always presented with the fastest speed first, and half with the slowest.

*Procedure.* Fifteen practice trials covering all the conditions were given on a day preceding the day of the experiment. In addition six practice trials covering each of the four sizes of window were given just before the experiment. The practice period, and the combined practice and experimental period, each took about 90 min.



*Results and discussion*

The results are shown in Table 3. The proportion of omissions ranged from about one-fifth at the lowest error rate to three-quarters at the highest. Owing to the relatively large individual differences, the mean errors for the moving windows of 79 and 9 letterspaces at reading speeds of 300 and 210 words per minute are not significantly larger than the comparable mean errors in Table 2, where about the same average rates of reading were used, although the standard errors are on average about twice as large. (In the present experiment all the subjects read at the same fixed speeds, whereas in Exp. I a separate speed was selected for each subject in order to give a fixed proportion of errors with the largest window. Thus, whereas in the present experiment the individual differences appeared entirely as differences in the number of errors, in Exp. I they appeared partly as differences in the rate of reading.)

Table 3. *Errors per 100 words in Exp. II*

No. of words exposed	Rate of reading (words per minute)									
	300		240		210		130		90	
	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
13.2	18.0*	3.3	9.0†	2.6	6.0*	1.5	—	—	—	—
1.8	29.8*	4.9	17.2†	4.2	11.5*	2.6	—	—	—	—
1.0	—	—	30.6	5.5	21.7†	4.7	4.9	1.0	—	—
0.6	—	—	—	—	46.2	4.9	20.1†	5.2	11.6	2.4

\* Not different from the comparable mean in Table 2.

† Not different from the comparable mean in Table 4.

The experimental conditions with the two smallest windows are plotted as crosses in Fig. 2; beside each cross is printed the mean number of errors per 100 words for that condition. Equally spaced straight lines passing through the origin have been fitted by eye to represent the error contours for 10, 20, 30 and 40 errors per 100 words. The contours, which are consistent with the data within the limits of the variability between subjects, represent exposure times for each letter of 0.38, 0.28, 0.22 and 0.17 sec., respectively.

The error contours suggest that for small windows the accuracy of reading was determined very largely by the time for which each letter could be seen; it did not matter whether the time resulted from a small window moving relatively slowly, or from a rather larger window moving more quickly. This finding is compatible with both the single-shot and the ciné camera models. The difference in prediction between the models concerns the exposure time at which the steep rise in errors occurred. Fig. 2 shows that the errors increased from about 20 to 40% as the exposure time of each letter was reduced from 0.28 to 0.17 sec. Even with the shortest exposure time all the letters could still be seen clearly, and there was no apparent reduction in the brightness of, or contrast in, the display. But as the observed error rate of 46% indicates, it was not possible to read very adequately. This favours the model of the visual system as a single-shot camera with an upper limit of about five shots per second, since reading should start to break down as the exposure time approaches the interval between shots.

The results for the moving windows of 9 and 79 letterspaces are not included in Fig. 2 because with larger windows the error rate was determined more by the speed of reading than by the size of the window. As the window increased beyond the range shown in Fig. 2 the error contours turn upwards, and finally run vertically.

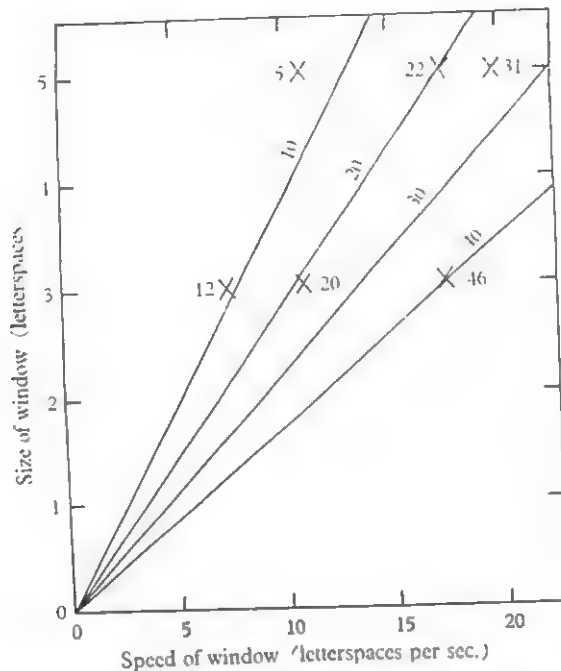


Fig. 2. Equal-error contours showing the relationship of speed of reading to size of window for a group of twelve subjects. Each cross marks an experimental condition. The value of the mean error per 100 words for this condition is printed beside the cross. The equally spaced contours have been fitted by eye.

### EXPERIMENT III. PRACTICE

A third experiment was designed to answer the questions: (a) can reading through small windows be improved by practice? (b) If so, what is the nature of the improvement?

#### Method

The twelve new experimental subjects were naval ratings aged between 18 and 27 years. A thirteenth rating developed a moderately severe stammer under the stress of the experiment, and had to be discarded.

The design is indicated in Table 4. Seven periods of practice, or combined practice and experiment, were spread over a week. Experimental tests were given at the end of the first period, and at the end of the last three periods. The experimental conditions were always presented in order of size of window, starting with no moving window and ending with the window of three letterspaces. In the first test six of the subjects all read the same six test passages, while the other six subjects read another six test passages. In each case a Latin square design was used.

The last three experimental tests each comprised only four conditions, the subject reading two new test passages and two which he had read in the first test. The order of the new and old passages was balanced over subjects, so that all twelve passages were read once in each condition of each of the last three tests. In each condition half the passages were read for the first time, and the other half were read for the second time.

Each of the seven practice periods comprised nine trials and covered all the experimental conditions which were to follow. The first period of combined practice and experiment took about 60 min. Subsequent practice periods took about 30 min., and the last three periods of combined practice and experiment took about 40 min.

## Results and discussion

The results are shown in Table 4. Where the window sizes and rates of reading are comparable, the means for the first period are not significantly different from the means in Table 3. In the tests during periods 5, 6 and 7 the passages which had been read once before give 6% fewer errors. Since this is too small a difference to be significant statistically, the results from both the old and new passages have been combined. Most of the improvement with practice occurred with the two smallest windows. The only statistically reliable evidence of continued improvement between periods 5 and 7 comes from the smallest window of three letterspaces; there was a reduction in all three kinds of error, but it was significant only for omissions ( $P < 0.05$ ) and insertions ( $P < 0.01$ ).

Table 4. Errors per 100 words in Exp. III

No. of words exposed	Rate of reading (words per minute)	Period no.							
		1		5		6		7	
		Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
13.2 (no moving window)	240	8.1	1.6	5.7	0.8	4.8	1.2	7.1	1.3
13.2	240	8.7*	1.9	—	—	—	—	—	—
2.6	240	12.8	2.2	13.2	2.9	9.8	2.1	9.2	1.6
1.8	240	13.5*	2.3	10.7	2.1	11.6	2.2	9.9	1.7
1.0	210	17.6*	2.5	8.0†	0.9	7.0	1.4	5.4†	0.8
0.6	120	16.0*	1.9	—	—	—	—	—	—

\* Not different from the comparable mean in Table 3.

† Practice 5—practice 7  $P < 0.01$ .

At the rates of movement of the two smallest windows each letter was visible for only about 0.3 sec. On the model of a single-shot camera with a rate of firing of not more than about five times per second, the reduction in omissions with practice would be ascribed to learning to read with very little overlap between successive views taken by the camera. The reduction in insertions is related to the observation that towards the end of the experiment the subject tended to wait at least until a whole word had been seen before saying anything, rather than guessing at the word sooner and subsequently correcting himself. Guessing would have been a less risky strategy if the syntactical structure of Basic English had been more familiar.

## EXPERIMENT IV. EYE MOVEMENTS

A final experiment was designed to show the effect of small windows upon the frequency of fixations and regressions.

## Method

The six new experimental subjects were again young naval ratings. Two additional ratings had to be discarded because they could not read aloud sufficiently quickly; another was discarded owing to excessive drift in his electro-oculographic potential.



*Apparatus.* The reading pacer was placed inside a screened box. The method of recording eye movements followed Shackel, Sloan & Warr (1958). Two electrodes embedded in small rubber cups (Shackel, 1958) were filled with electrode jelly and attached by suction one just beyond the outer corner of each eye. In some cases a small pit was first made in the epidermis with a dental burr (Shackel, 1959). An earthing electrode was attached below and behind one ear. The changes in electrical potential developed across the eyes were fed through a direct-current E.M.I. amplifier to an Evershed and Vignoles pen recorder.

The *design* and *procedure* duplicated principally the prose part of Expt. I.1. After the practice the electrodes were attached, and the subject's head was clamped in a comfortable position. The tape recorder had to be switched off, since it was in an unscreened box, but the subject still had to read aloud. The experimenter deleted any parts of the pen record made while the subject was omitting to speak a phrase or line. The sizes of window were presented in a different random order to each subject. The mean rate of reading was 260 words per minute, s.e. between subjects 9.3. At the end the window showing only 3 letterspaces was presented at a mean speed of 115 words per minute, s.e. 10.9. The practice and recording together took up to 2½ hr.

Two of the subjects subsequently practised for five periods, as in Exp. III. In the seventh period their eye movements were recorded again under the same conditions as in the first period.

*Scoring* was restricted to the middle part of each line of record, since the largest windows were only filled with letters here. An average of about 15 sec. of record was selected from each subject under each condition, and the number of fixations and regressions during this period of time was counted. Some of the records contained spikes produced by action potentials in the facial muscles while speaking. These should have affected to about the same extent the records obtained with all sizes of window except the smallest, since the rate of reading was held constant over this range. Thus they may bias the mean number of fixations and regressions, but should not affect the comparisons between sizes of window, except between the smallest window and the remainder.

### *Results and discussion*

*Fixations* averaged 3.8 per second, s.e. between subjects 0.12. There were no obvious differences between reading with any of the sizes of window, and the frequency of fixations remained about the same after seven practices with small windows ( $P > 0.05$ ).

The failure of fixations to rise above five per second in frequency as the size of the window was reduced is compatible with the model of the visual system in reading as a single-shot camera having a limit to its rate of firing of about five shots per second; for there would be no point in the eyes fixating the line more frequently than the camera could take views. Perhaps the rate of fixations of good readers remains fairly constant under different conditions of reading (Woodworth, 1938, p. 724) because it is set to match the maximum comfortable frequency at which the central computer can accept successive blocks of information.

For windows exposing five or more words at a time *regressions* averaged 3.0% of all fixations, and involved four of the six subjects. For windows moving at the same speed and exposing less than three words at a time, regressions averaged 1.2% and involved two of the subjects. This difference is not reliable statistically. With the smallest window which moved more slowly, regressions averaged 2.9%.

No obvious advantage was to be gained from a regression when reading through a very small window, since the subject could not see any typescript when his eyes regressed. Perhaps a regression simply indicated a decision not to take in any more information until he had thought about what he had just seen. It is possible that the window was the direct cause of some of the regressions, the subject looking away from the display for a brief period when he had lost track of what he was supposed to be reading. These suggestions are in line with the view that regressions are largely the

result of difficulty in reading, not the cause. Clearly, making the student read through a window which moves continuously, as the Carborundum (1959) films do, is unlikely to prevent regressions.

The experiments were carried out under the general direction of Dr N. H. Mackworth, who lent the eye-movement recorder. Dr M. Stone and Miss V. R. Cane gave advice on statistical methods, and Dr R. Conrad and Mr B. Shackel commented on a draft manuscript. The experimental subjects were supplied by the Royal Navy. Financial support from the Medical Research Council is also gratefully acknowledged.

## REFERENCES

- ANDERSON, I. H. & SWANSON, D. E. (1937). Common factors in eye-movements in silent and oral reading. *Psychol. Monog.* **48**, no. 3, 61-9.
- CARBORUNDUM CO., LTD. (1959). *An Experiment in Industrial Films*. Manchester.
- EDWARDS, A. L. (1951). Balanced Latin-square designs in psychological research. *Amer. J. Psychol.* **64**, 598-603.
- FLESH, R. (1948). A new readability yardstick. *J. Appl. Psychol.* **32**, 221-33.
- LUCKIESH, M. & MOSS, F. K. (1942). *Reading as a Visual Task*. New York: Van Nostrand.
- PERRY, W. G., JR & WHITLOCK, C. P. (1948). *Harvard University Reading Films Series Two*. Cambridge, Mass.: Harvard University Press.
- PERRY, W. G., JR & WHITLOCK, C. P. (1954). A clinical rationale for a reading film. *Harvard Educ. Rev.* **24**, 6-27.
- PIERCE, J. R. & KARLIN, J. E. (1957). Reading rates and the information rate of a human channel. *Bell System Tech. J.* **36**, 497-516.
- POULTON, E. C. (1950). Perceptual anticipation and reaction time. *Quart. J. Exp. Psychol.* **2**, 99-112.
- POULTON, E. C. (1961). British courses for adults on effective reading. *Brit. J. Educ. Psychol.* **31**, 128-37.
- SHACKEL, B. (1958). A rubber suction cup surface electrode with high electrical stability. *J. Appl. Physiol.* **13**, 153-8.
- SHACKEL, B. (1959). Skin drilling: a method of diminishing galvanic skin-potentials. *Amer. J. Psychol.* **72**, 114-21.
- SHACKEL, B., SLOAN, R. C. & WARR, H. J. J. (1958). Detector plots eye movements. *Electronics*, 31 Jan.
- SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.
- TINKER, M. A. (1943). Illumination intensities for reading newspaper type. *J. Educ. Psychol.* **34**, 247-50.
- WELFORD, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quart. J. Exp. Psychol.* **11**, 193-210.
- WELLS, H. G. (1951). *The Time Machine*. Put into Basic English by K. Robinson. London: Basic English Publishing Co.
- WOODWORTH, R. S. (1938). *Experimental Psychology*. New York: Holt.

(Manuscript received 28 November 1961)





## ON THE PSYCHOLOGICAL NATURE OF STAGE IMPERSONATION

By R. NATADZE

Tbilisi, U.S.S.R.

The hypothesis examined in the present paper is that it is *fixated set* evolved on the basis of picturing to oneself the particular imaginary situation imposed by the play that constitutes the foundation on which stage impersonation rests. Findings are presented of some experiments on the ability to evoke a set on the basis of an imagined situation. The subjects in these experiments were grouped as follows: (a) gifted actors and promising students of the Tbilisi Theatrical Institute; (b) persons unconnected with the theatre; and (c) persons incapable of stage impersonation. The data suggest a high correlation between the ability to evolve a set on the basis of imagination and the capacity for stage impersonation.

### I. INTRODUCTION

When, with a spontaneity convincing to the spectators, the actor acts in conformity not with a real, but with an *imagined* situation imposed upon him by the play, embodying thus in his stage performance experiences and conduct not his own—what is the psychological mechanism of such behaviour?

This constitutes the essence of the time-honoured problem of the psychological bases of stage impersonation which for centuries has been discussed as the moot issue of 'genuine mental experience' or 'real feeling' *versus* 'cold virtuosity' and 'imitation'.

In the opinion of the writer, the stagnation which has up to now marked discussion of problems implied in the above controversy is due to the fact that in recent times the subject has, as a rule, been discussed by theatrical workers, without reference to the evidence provided by present-day psychological science.

As a result of systematic observation of the preparation of roles by the students of the Tbilisi Theatrical Institute, of attending their rehearsals, and of a psychological analysis of the theory of stage performance known as the Stanislavsky system, (Stanislavsky, 1938), the writer has been led to the conviction that stage behaviour attains convincing naturalness, i.e. 'impersonation' is attained, only when the actor succeeds in evolving and fixating a set corresponding to the *imagined* situation called for by the play. This situation once fixated, i.e. having become a 'fixated set' (Uznadze, 1939), conditions corresponding behaviour on the part of the actor in much the same way as real-life behaviour is immediately conditioned by a corresponding set emerging on the basis of the subject's perception of the real situation.

But can such a set emerge on the basis of an *imagined* situation and under conditions of the subject's *awareness of its unreality*? The writer thinks himself justified in asserting that this possibility has been proved experimentally even for perceptual situations the content of which is at variance with that imagined (Natadze, 1960).

If the assumption is correct, that it is the corresponding set of the actor that constitutes the actual basis of the 'mechanism' of stage impersonation, then the ability

to induce in oneself an imaginably based set, especially when there is contradiction between the imagined and the perceived, should correlate with the capacity for stage impersonation.

## II. THE EXPERIMENTS

The tests were carried out with the following groups of subjects:

(1) Ten prominent dramatic actors: eight from three Tbilisi theatres (two Georgian and one Russian) and two former actors of the Moscow Art Theatre.

(2) Seven younger actors (also hailing from Tbilisi) with two or three roles to their credit, created in the course of a few years' stage work.

(3) Ten promising students of the Georgian Theatrical Institute.

(4) Seven average students of the same Institute.

These four groups of subjects were classed as belonging in the category of actors capable of stage impersonation ('actors' in the Tables) the first three groups consisting of persons with clearly defined abilities in this respect.

(5) Four students of the Theatrical Institute, who had proved unsatisfactory and had therefore been advised to leave the Institute.

(6) A prominent Georgian comedian (hereinafter referred to as 'N') known for his incapacity for impersonation: he invariably displays one, and only one, mode of comic acting.

The latter five subjects were selected as persons definitely devoid of the ability for stage impersonation.

(7) Twenty-one subjects not connected with the theatre. Of these, thirteen were research psychologists and eight clerical and administrative workers. This group was taken as belonging to the category of subjects who had nothing to do with stage activity. The possibility is of course not to be excluded that some of them had potential acting ability. One reason for singling out psychologists was that as a result of their professional habits (e.g. introspection) they, unlike other persons not connected with the theatre, are accustomed to dealing with mental imagery. (Groups 5-7 are classed together in the Tables as 'non-actors'.)

Since stage impersonation is mainly characterized by the actor having to act conformably to an imagined situation while actually perceiving an entirely different one (partners, wings, sets, etc.), we shall deal with the results only of the experiments in which the imagined is contradicted by the perceived. The experimental method used has been described elsewhere (Natadze, 1960).

*Experiment 1.* The subject is instructed to inspect two equal-sized wooden balls (with handles) and imagine as vividly as possible that he is simultaneously lifting them by the handles. At the same time he is to imagine that one of the balls is very heavy and the other very light. When he has succeeded in doing this—i.e. in going through the 'experience' of 'feeling' the difference, he has to go through fifteen successive rehearsals of the imagined procedure. This is followed by the critical test. With his eyes closed, the subject lifts two objects of equal weight by their handles and compares them as to weight. If his perception is distorted (i.e. if one of the objects is perceived by him as being heavier), this is an indication of the effect on his perception of the fixated set evolved on the basis of the imagined lifting of unequal weights; viz. a set for perceiving a heavier object on one side as compared to the other. If no illusion occurs, i.e. if in the critical test the objects are, from the very start, perceived as being equal, it is an indication that a fixated set based on lifting *imagined* weights has not been evolved. The earlier the subject, in the course of the critical tests, shifts to adequate perception, the fainter is the effect of the fixated set.

The results of Expt. 1 are given in Table 1.

It is evident from Table 1 that the percentage of set fixation under the above-described conditions is much higher for groups more or less capable of stage impersona-

tion than for subjects unconnected with the theatre, though this latter group includes psychologists. Of the eight non-psychologists tested in this experiment, only one developed a set-induced illusion. Of the seventeen actors and actresses only one failed to evoke an illusion on the basis of imagination. This was subject A, a comedian. Of the ten promising students of the Theatrical Institute, two gave negative results; one of them (here referred to as B) is a well-defined comedian, while the other (a girl student L, at present playing minor roles) is not considered to be an actress of high promise.

Table 1. *Experiment 1*

Subjects	N	Adequate perception	Total illusions
Prominent actors	10	1	9
Young actors	7	6	7
Promising students	10	2	8
Average students	8	2	6
Total 'actors'	35	5	30
Psychologists	13	5	8
Clerical and administrative	8	6	2
Unsatisfactory students	4	3	1
Comedian N	1	1	0
Total 'non-actors'	26	15	11

Table 2. *Experiment 2*

Subjects	N	Charpentier illusion	Disturbance of Charp. illusion	Adequate perception	Perception of contrast
Prominent actors	10	2	8	3	5
Young actors	7	1	6	2	4
Promising students	10	1	9	3	6
Average students	6	2	4	3	1
Total 'actors'	33	6	27	11	16
Psychologists	7	5	2	1	1
Clerical and administrative	8	6	2	2	0
Unsatisfactory students	4	3	1	0	1
Comedian N	1	1	0	0	0
Total 'non-actors'	20	15	5	3	2

It should also be noted that of the four unsatisfactory students only one developed a set-induced illusion. To this it ought to be added that *not one of the tests* to which N was subjected resulted in an illusion. Mention should also be made of the fact that subjects capable of impersonation not infrequently develop a *stable* illusion, i.e. they manifest the effects of intensive and prolonged action upon them of imagination-stimulated set, while in the rest of the subjects a stable set occurs but seldom.

*Experiment 2.* The overcoming of the Charpentier illusion on the basis of a fixated set stimulated by imagination. The subject, while visually concentrating on Charpentier cylinders, is to picture to himself as vividly as possible that he is lifting them by their strings, and that in so doing he feels the smaller cylinder to be very heavy and the larger one very light (as if it were hollow). After fifteen successive rehearsals of this imagined procedure



the subject, in the critical trial, lifts the Charpentier cylinders, but now his task is actually to compare them as to weight.

The results are given in Table 2.

Of the group of prominent actors, two gave no evidence of the Charpentier illusion having been disturbed, i.e. of having developed a fixated set: the comedian A of Expt. 1, and also one other subject; and of the group of capable students, the comedian B, referred to above. The fairly large percentage of illusions in the group of psychologists may partly be due to their having had much experience with the Charpentier cylinders.

*Experiment 3.* Induction of a set for imagined lifting of *unequal weights* while in the act of *actually lifting equal loads*.

After the subject has simultaneously lifted, by the handles, two wooden balls of equal size and of equal weight, and has thereby made sure of their being of equal weight, he is told to lift these same objects, each time forcing himself into 'feeling', 'experiencing', as clearly as possible, that one of them (say, the right-hand one) is very heavy, and the other, very light. After this task has been gone through fifteen times, the subject is instructed in the 'critical' test to shut his eyes and to lift by the strings two objects (say, wooden cylinders) of equal weight, and compare them.

The results of the experiment are given in Table 3.

Table 3. *Experiment 3*

Subjects	N	Adequate perception	Total illusions
Prominent actors	10	0	10
Young actors	7	1	6
Promising students	10	0	10
Average students	7	2	5
Total 'actors'	34	3	31
Psychologists	13	6	7
Clerical and administrative	8	5	3
Unsatisfactory students	4	3	1
Comedian N	1	1	0
Total 'non-actors'	26	15	11

As is seen from the table, in two of the groups, that of prominent actors and that of promising students, set-induced illusions were produced in all the subjects without exception; of the younger actors, only one (C, at his best in comic parts) showed a negative result.

It should be noted that for subjects not connected with the theatre the task in this experiment proved to be much more difficult than the tasks in the foregoing experiments. The majority of the subjects managed to attain the required clear-cut kinaesthetic visualization of lifting unequal weights only at the cost of great exertion and after no less than ten to sixteen vain attempts; this may be accounted for by the strongly inhibitive action of the objectively equal weight of the objects lifted. With a considerable number of the subjects of this category no sufficiently clear impression could be produced, whatever the exertion made to that end. In this experiment the difficulty apparently consists in that *within one and the same modality* the

imagined content is contradicted by the perceived content (kinaesthetic visualization of unequal weights contradicted by actual perception—also kinaesthetic—of equal weights).

It is significant that actor subjects achieved the required effect with extreme ease, and usually at the first or second trial. One would infer from the foregoing that for them the task in this experiment proved to be still easier than those of the preceding experiments. Presumably the possibilities for 'acting' inherent in this experiment serve to facilitate 'performing' on the part of these subjects: the subject is actually lifting objects, and this should make it the easier for him to act, while engaged in the process, the part of one lifting unequal weights: accordingly, he puts more effort into the muscles of one hand, slows down the movement of that hand, applies a tighter grip to the string of one of the two objects, etc. In the first two experiments such a possibility is ruled out, since the subject's hands lie motionless on the table, in a position that has no relevance to the task. The actor's habit of working out, on the stage, a relevant attitude toward what is being imagined must, in all probability, have contributed its share.

*Experiment 4.* Induction of a fixated set on the basis of an imagined lifting of unequal weights while the subject is actually lifting loads whose weight correlation is contradicted by that mentally pictured.

The subject simultaneously lifts, by their strings, two wooden cylinders, of 123 and 186 g., respectively, to make sure that one of them, say the right-hand one, is heavier than the other. He is then told to lift the same objects by the strings and at the same time force himself into 'experiencing' as vividly as possible (as if it were a sensation) the sensation that the objectively lighter cylinder is very heavy. After the subject has succeeded in doing this consecutively fifteen times, he is, with his eyes shut, put to the critical test, which consists in his having to lift simultaneously two equally heavy objects, and compare them as to weight.

The results of this experiment are given in Table 4.

Table 4. *Experiment 4*

Subjects	N	Adequate perception	Total illusions
Prominent actors	10	0	10
Young actors	7	1	6
Promising students	10	0	10
Average students	7	2	5
Total 'actors'	34	3	31
Psychologists	13	5	8
Clerical and administrative	8	8	0
Unsatisfactory students	4	3	1
Comedian N	1	1	0
Total 'non-actors'	26	17	9

Although the numerical data of this experiment almost coincide with those of the preceding test (barring the non-psychologists unconnected with the theatre, who failed to produce a set-induced illusion in this test), in the latter case the *difficulty* experienced by subjects unconnected with the theatre in performing the task increased, whereas actors achieved it as readily as in the preceding experiment.

*Experiment 5.* Induction of a fixated set on the basis of a haptically imagined unequal volume, while the subject is actually engaged in perceiving haptically an equal volume.

The subject holds equal balls—one in each hand—grasping them tightly with his fingers. He is to visualize as vividly as possible (as if it were an actual 'experience') that he is grasping a large ball with one hand (the right, say), and a quite tiny one with the other. After fifteen consecutive successful rehearsals of the procedure required, the subject with his eyes closed, has, in the critical test, to compare as to volume two equal wooden balls which the experimenter places momentarily in his hands.

The results are presented in Table 5.

Table 5. *Experiment 5*

Subjects	N	Adequate perception	Total illusions
Prominent actors	10	2	8
Young actors	7	2	5
Promising students	10	2	8
Average students	6	2	4
Total 'actors'	33	8	25
Psychologists	13	8	5
Clerical and administrative	8	8	0
Unsatisfactory students	4	4	0
Comedian N	1	1	0
Total 'non-actors'	26	21	5

The task set in this experiment proved to be the most difficult one. Even actors attained the required result with some difficulty, while most of the subjects unconnected with the theatre who succeeded in carrying out the instructions, did so by dint of utmost exertion, not infrequently manifesting great fatigue towards the end of the test.

This difficulty is presumably due, first, to the circumstance that the instruction to grip the balls tightly precludes the possibility of 'acting', i.e. of making movements conformably to what is being imagined; and secondly, to the tactile-haptic perception being linked with a markedly pronounced 'feeling for reality' which apparently causes a special difficulty in overcoming this sensory impression. As is seen from the table, the percentage of fixated set showed a considerable drop in this experiment.

Incidentally, among those who again failed to evoke a set were the comedian N, and the girl student L (mentioned under Expt. 1, above).

*Experiment 6.* Appraisal of the effect on the perception of critical objects of a set based on imagination.

The subject has to look at two equal-sized wooden cylinders and imagine that he is lifting them by strings attached to them. At the same time he should imagine that one of them is very heavy, and the other very light. After fifteen rehearsals of this imagined handling of the two cylinders, the subject has, in the critical test, momentarily to lift them by their strings, in order to compare their real weights, which are respectively 132 and 186 g. This difference is considerably in excess of that of the threshold: for this reason the subject, as a rule, unmistakably perceives the



difference in weight without any preliminary set-inducing trials. The conditions of the experiment make it possible to take into account only the set's *contrastive* effect: it was the objectively heavier cylinder that was conceived by the subject as being the heavier one.

The results of the experiment are presented in Table 6.

Table 6. *Experiment 6*

Subjects	N	Adequate perception	Total illusions	Illusion of equality	Illusion of contrast
Prominent actors	10	2	8	1	7
Young actors	7	3	4	1	3
Promising students	10	2	8	2	6
Average students	6	2	4	1	3
Total 'actors'	33	9	24	5	19
Psychologists	13	11	2	1	1
Clerical and administrative	8	8	0	0	0
Unsatisfactory students	4	4	0	0	0
Comedian X	1	1	0	0	0
Total 'non-actors'	26	24	2	1	1

As is seen from Table 6, it was the potency of the effect which imaginally-based fixated set had on perception that proved to be the most characteristic, the most specific manifestation of such a set in persons capable of stage impersonation. In an overwhelming majority of subjects in this category, the imaginally stimulated fixated set has such a potent effect on the subsequent perception of weights that it tends to blot out differences amounting to no less than 41% of the weight of the lighter object. This effect is shown by the table to differ in degree: when the contrastive effect is strong, an illusion of a *reversed* weight relationship emerges, so that the heavier object is perceived as being the lighter one; when the contrastive effect is less strong, an illusion of equality is the result. Incidentally, with persons capable of stage impersonation the first variety of illusion, i.e. that with the higher potency of effect, predominates.

It is worth noting that of the subjects unconnected with the theatre, only two psychologists produced a positive result. Both are of the type in whose inner life imagination plays a great part. In all of the preceding tests they readily performed the task assigned them and produced a strong and stable illusion. They constitute just under 10% of all the subjects unconnected with the theatre, whereas 80% of the prominent actors and promising students and 74% of all the actors and students, including those of average ability, give a positive result.

### III. DISCUSSION

1. The objective data obtained unequivocally suggest that the ability to evoke a fixated set based on an imagined situation is far more widespread among persons capable of stage impersonation than among those unconnected with the theatre: again, in persons incapable of impersonation this ability is, as a rule, absent, since of the five subjects of the latter category tested in the present series of experiments (four students and one actor, N), only one student displayed this ability, and that, too, only in the first four tests.

2. Three of the actors and promising students tested, who in a number of trials failed to develop a strong fixated set, were comedians (the subjects A, B and C). Thus, all the four comedians (including the above-mentioned N) tested showed a low capacity for evolving a set based on imagination. The question arises whether this is not characteristic of comedians generally. Does not the comical nature of the roles performed by them prevent genuine impersonation? For does not an actor performing such a role exhibit, as it were, to the public his critical attitude towards the character he personifies?

Be that as it may, the following observation is worth recording: had the comedians not been included in the group of 'actors', in most of our experiments a set-induced illusion would have been produced by all the nine prominent actors and by 90% of the young actors and promising students. The student L is the only exception: in a number of experiments she produced no illusion at all. It should be observed, however, that later her abilities in stage performance proved to be of no high order.

3. Of all the experiments, the most striking difference between the two main categories of subjects was revealed in the experiment (Expt. 6) designed to ascertain the effect on perception of the imaginably stimulated set.

Second in this respect are those experiments (Expts. 3-5) in which an imaginably stimulated set of a given modality emerges in despite of actual perception in the same modality (Expts. 3-5), i.e. under conditions of maximum contradiction between the imagined and the perceived situation, which, as has above been observed, should be characteristic of the actor.

4. The difference between actors and subjects unconnected with the theatre, as shown in the degree of difficulty with which they voluntarily developed a set with respect to an imagined situation, proved to be particularly pronounced; especially so in those experiments in which the imagined situation is contradictory to that perceived (Expts. 3-5). All the actors tested attained, with great facility and at once, the specifically active attitude toward the imagined situation which is necessary for evolving such a set, whereas the subjects having no connexion with the theatre achieved this with great difficulty, by dint of a strong inner strain. The former can easily 'experience' or 'feel' the imagined, while those of the latter who can attain such a result achieve it only with great difficulty and at the cost of very considerable inner tension.

5. It should be specially emphasized that subjects who have sufficiently familiarized themselves with the experimental procedures can easily discriminate between a mere 'conceiving' of what they are to imagine and a specific 'inner attunement' towards the imagined. The actor subjects, as well as the psychologist subjects, often refer to this latter: 'I've worked myself into the (mental) state', 'I'm in the (mental) state'; and almost invariably make use of such terms as 'preparedness', 'directedness', etc.

This cannot but bring to mind analogous observations by Stanislavsky. According to him, the main and indispensable condition for convincing behaviour on the stage is 'belief' in the reality of what is being acted, though this, of course, should not imply any actual lessening of the actor's sense of the real.

In the present writer's opinion, the inner state of the actor to which Stanislavsky

refers as 'belief' in the reality of what is being acted on the stage, is just that real inner state of specific active attunement to the imagined which constitutes the necessary prerequisite for evolving a set with respect to the imagined, the subject being at the same time fully aware of its unreality.

6. A strong argument in favour of attitude as a factor in the actor's stage behaviour would seem to be the experimentally established fact that an imaginably stimulated fixated set may, while functioning against the background of a general attitude, act upon the actual reality with no impairment of the 'sense of the real' resulting therefrom, i.e. the actor being fully aware of the unreality of the imagined situation: that is why the actor's *acting* does not change into virtual *behaving*. This would be impossible if the actor had been acting under the influence of a genuine and earnest feeling evoked by an imagined situation.

#### REFERENCES

- NATADZE, R. (1960). Emergence of set on the basis of imaginal situations. *Brit. J. Psychol.* **51**, 237-45.  
STANISLAVSKY, K. (1938). *Rabota aktyora nad soboy* (The Actor's Work on Himself). 2nd ed., Moscow.  
USNADZE D. (1939). Untersuchungen zur Psychologie der Einstellung. *Acta Psychologica*, **4**, 323-60.

*(Manuscript received 24 January 1961)*





## BRIGHTNESS JUDGEMENTS AND STIMULUS SIZE AND DISTANCE

BY RAY OVER

*Psychology Department, Queen's University, Belfast*

Three groups of subjects were instructed to judge the relative brightness of two stimuli, the sizes and distances of which were varied to allow three visual angle relationships. Testing was carried out under reduced viewing conditions and the groups differed in terms of the criteria subjects were instructed to use in making brightness judgements. For all groups it was found that when the two stimuli were equal in luminous intensity the stimulus subtending the smaller visual angle was consistently judged to be less bright than the stimulus subtending the larger visual angle. Judgements made with 'projective' criteria were more a function of stimulus size and distance than were judgements made with either 'objective' or 'apparent' criteria.

### I. INTRODUCTION

It will be asked whether: (a) two stimuli of the same size but at different distances from the subject, and (b) two stimuli of different sizes but at the same distance from the subject, will be judged to be equal in brightness when they are of the same luminous intensity. Gibson (1950, footnote, p. 137) has denied that brightness judgements are a function of stimulus distance, stating, '... the presumable assumption (is) that an object necessarily appears darker as its distance from the eye increases. Apart from the little-known effects of atmospheric conditions on visibility—aerial perspective—the assumption has no basis in physical fact.' There are, however, two retinal configurations which are a function of the luminous intensity of the stimulus, and one of these is also a function of stimulus size and distance (Ittelson & Kilpatrick, 1961). The total amount of light reaching the nodal point of the eye from a stimulus of constant luminous intensity is an inverse square function of distance and a direct function of stimulus area. The solid angle subtended by the stimulus is also an inverse square function of distance and a direct function of stimulus area. Accordingly, the intensity of light subtended (the amount of light per unit solid angle) is independent of stimulus area and distance, whereas the total amount of light subtended is a function of stimulus area and distance.

Are there any experimental conditions under which brightness judgements are affected by variation in stimulus size and distance? Some size and distance discrimination data suggest there are such experimental conditions. Both size-constancy (Gilinsky, 1955) and distance-constancy (Over, 1961) relationships are found under unrestricted viewing conditions when subjects are instructed to use 'objective' judgemental criteria. Under these viewing conditions subjects given 'projective' instructions judge two stimuli to be equal in size only when they subtend equal visual angles (Gilinsky, 1955), while the size judgement of subjects instructed to judge in terms of 'apparent' size follow a relationship somewhere between the relationships found for 'objective' and 'projective' instructions (Thouless, 1931). Under reduced viewing conditions (darkroom, monocular viewing, restriction of head-movements, etc.), however, two stimuli are judged to be equal in both size and dis-

tance only when they subtend equal visual angles (Over, 1960), irrespective of whether the subject is instructed to use 'objective' or 'projective' criteria in making judgements (Over, 1960*a*).

Under reduced viewing conditions subjects given 'objective' judgemental criteria judge two objects of different luminous intensities to be equal in both size (Robinson, 1954) and distance (Coules, 1955), when the stimulus of greater luminous intensity subtends a smaller visual angle than the stimulus of lesser luminous intensity. Stimulus size, distance, and luminous intensity are not, however, equivalent operations for size and distance judgements. Although some limited changes in size and distance judgements can be brought about by variation in any of stimulus size, distance, and luminous intensity, two stimuli are not judged to be equal in size and distance only when they subtend equal amounts of light (except when they also subtend equal visual angles).

If variations in the sizes and distances of stimuli affect brightness judgements in the manner that variation in the luminous intensities of the stimuli affects size and distance judgements, it would be expected that under reduced viewing conditions two stimuli would be judged to be equal in brightness when a stimulus subtending a smaller visual angle is of a greater luminous intensity than a stimulus subtending a larger visual angle. When subjects are instructed to use 'projective' judgemental criteria it would be expected that two objects would be judged to be equal in brightness when they subtended the same total amounts of light. Even when subjects are instructed to use 'objective' criteria, that is, when they are instructed to judge that two objects are equal in brightness when they subtend the same light per unit solid angle, it would be expected that brightness judgements would be, to some extent, a function of stimulus size and distance. If an instruction to judge in terms of 'apparent' brightness is as ambiguous as an instruction to judge in terms of 'apparent' size, it would be expected that brightness judgements would be more a function of stimulus size and distance when the subject is instructed to judge 'apparent' brightness than when the subject is instructed to judge 'objective' brightness.

The sole experimental evidence is provided by Hanes (1951) who found that, for certain visual angle differences and luminance levels, the larger of two stimuli which were at the same distance from the subject had to be of the lesser luminous intensity before the two stimuli were judged to be equal in brightness. It is not known whether Hanes's experimental situation was reduced (in terms of size and distance discrimination functions): nor does Hanes set out explicitly the judgemental criteria his subjects were instructed to use. It is possible that Hanes's results are artificial: Hanes has not demonstrated that even when subjects are explicitly instructed to make judgements independent of stimulus size and distance brightness judgements under reduced viewing conditions are to some extent a function of stimulus size. Nor has Hanes examined the situation where brightness judgements are made of two stimuli of the same size but at different distances from the subject. The experiment to be reported is aimed at providing these further data.



## II. PROCEDURE

The reduced viewing conditions used were similar to those set out elsewhere (Over, 1962). Under these conditions it has been found that two stimuli are judged to be equal in size and distance only when they subtend equal visual angles (Over, 1960). The testing was carried out in a darkened room and the subjects, tested individually, were seated inside a booth containing a headrest. Viewing was monocular with the right eye. When the curtain in front of the subject was raised the subject was able to see two diamond-shaped lights at eye-level in front of him. Each stimulus was produced by the illumination of a milk-glass screen. An iris mechanism between a 15 W., 230 V. lamp and the screen enabled the experimenter to change the luminous intensity of the stimulus without varying the temperature of the lamp filament. Before the experiment luminance values (the average of twenty readings), measured with an S.E.I. photometer, were obtained for various iris apertures with each stimulus. Stimulus size was varied by placing a sheet of black cardboard with a specified aperture in front of the milk-glass screen. The experimenter was able to place the stimuli, which were separated by a constant visual angle of  $9^\circ$ , at either the same or different distances from the subject.

Three groups, each of ten subjects, were used. Most of the subjects were second or third year psychology students and all reported that they understood their instructions. The groups differed in terms of the judgemental criteria subjects were instructed to use. Each subject was instructed: 'When the curtain is lifted you are to judge whether the object on your left is brighter or less bright than the object on your right.' In addition, the following specific instructions were given:

'Objective' instructions. 'You are to look at a unit area, say the square inch in the centre, of each object and to see whether the two objects have the same amount of light in that unit area. If the object on your left contains more light per unit area it is brighter than the object on your right. Do you understand that this criterion requires you to ignore overall differences in the sizes and distances of the objects?'

'Projective' instructions. 'You are to compare the two objects in terms of the total amount of light from each which reaches your eye. Do you understand that the total amount of light from an object which reaches your eye depends upon the size of the object and its distance from you? When two objects of different sizes emit light of the same intensity, that is, when they emit the same amount of light per unit area, the total amount of light reaching your eye is less for the smaller object than for the larger object.'

'Apparent' instructions. 'You are to judge whether the object on your left appears to you to be brighter or less bright than the object on your right.'

Table 1. Arrangements of stimulus sizes and distances

Stimulus size and distance combinations	Experimental arrangements								Ratios SS/SD; CS/CD
	SS	CS	SD	CD	SS	CS	SD	CD	
SS = CS SD = CD	2	2	7	7	4	4	14	14	1:1
SS = CS SD < CD	2	2	7	14	4	4	7	14	2:1
SS = CS SD > CD	2	2	14	7	4	4	14	7	1:2
SS < CS SD = CD	2	4	7	7	2	4	14	14	1:2
SS > CS SD = CD	4	2	7	7	4	2	14	14	2:1
SS ≠ CS SD ≠ CD	2	4	7	14	4	2	14	7	1:1

where SS/SD = CS/CD

SS signifies standard size in inches (diagonals); CS comparison size in inches; SD standard distance in feet; and CD comparison distance in feet.

The six combinations of stimulus size and distance, together with the twelve arrangements actually used in the experiment, are set out in Table 1. The seventh possible combination, where the two stimuli differ both in size and in distance but subtend unequal visual angles, was not used. The six combinations, for the stimulus values in Table 1, reduce to three different relationships between the visual angles subtended by the two stimuli.

The luminous intensity of the standard stimulus remained constant during the experiment at 2.80 log ft.-L. The luminous intensity of the comparison stimulus could be set by the experimenter at any one of twelve 0.01 log ft.-L. steps between 2.30 log ft.-L. and 1.40 log ft.-L. For each stimulus arrangement the subject made judgements when the luminance of the comparison stimulus was 2.30, 2.80, and 1.40 log ft.-L., the order of presentation of these luminances being varied at random between arrangements and subjects. If the subject judged 'less bright' when the luminance of the comparison stimulus was 2.30 log ft.-L. and 'brighter' when it was 2.80 log ft.-L., he was then required to make brightness judgements when the luminous intensity of the comparison stimulus was at each of the four intermediate levels, the order of these presentations varying at random between arrangements and subjects. If a judgement of 'brighter' was given when the comparison luminance was 1.40 and of 'less bright' when it was 2.80 log ft.-L., the subject made brightness judgements at the four intermediate luminance levels, the order of presentation again varying at random. Therefore, except when judgements of 'brighter' were given when the luminance of the comparison stimulus was 2.30 log ft.-L. and of 'less bright' when it was 1.40 log ft.-L., each subject made judgements at seven different luminous intensities of the comparison stimulus for the twelve stimulus arrangements.

### III. RESULTS

The cumulative frequency of 'brighter' judgements given by each instruction group to the three visual angle relationships of the stimuli with variation in the luminous intensity of the comparison stimulus are set out in Table 2. The luminous intensity of the standard stimulus was constant at 2.80 log ft.-L. The null hypothesis that as many judgements of 'brighter' as judgements of 'less bright' will be given when the two stimuli are of the same luminous intensity, irrespective of the visual-angle relationship of the two stimuli and of the instructions given to the subject, can be tested by a series of binomial tests of significance (Siegel, 1956, p. 40). The various  $z$  values, and the probability of obtaining these by chance alone, are set out in Table 3. One-tail probabilities are given for the two sets of cases where the stimuli subtend different visual angles, two-tail probabilities for the set of cases where the stimuli subtend equal visual angles.

The results show that when the two stimuli subtended equal visual angles at

Table 2. *Cumulative frequencies of 'brighter' judgements with change in the luminous intensity of the comparison stimulus*

Comparison luminances (log ft.-L.)	Visual angle relationships								
	2:1			1:1			1:2		
	P	A	O	P	A	O	P	A	O
1.4	—	—	—	—	—	—	29	—	—
1.3	—	—	—	—	—	—	19	40	—
1.2	—	—	—	—	—	—	13	36	40
1.1	—	—	—	40	40	40	5	26	37
1.0	—	40	40	38	39	37	1	13	26
2.9	—	39	38	34	34	30	—	9	14
2.8	40	31	33	25	21	20	—	1	6
2.7	39	25	19	—	—	—	—	—	1
2.6	35	14	11	4	3	2	—	—	—
2.5	32	1	2	2	—	—	—	—	—
2.4	24	—	—	2	—	—	—	—	—
2.3	11	—	—	—	—	—	—	—	—

P signifies 'projective', A 'apparent', and O 'objective' judgemental criteria.

the subject's eye as many judgements of 'brighter' as of 'less bright' were given when the stimuli were equal in luminous intensity. When, however, the comparison stimulus subtended a visual angle which was twice as large as the visual angle subtended by the standard stimulus, there were significantly more judgements of 'brighter' than of 'less bright' when the stimuli were of equal luminous intensities. When the visual angle subtended by the comparison stimulus was half the visual angle subtended by the standard stimulus significantly more judgements of 'less bright' than of 'brighter' were given when the stimuli were of equal luminous intensities. Judgements followed the above functions for each of the twelve stimulus arrangements set out in Table 1.

Table 3. *z* values and probabilities from the binomial tests

		Visual angle relationships								
		2:1			1:1			1:2		
		P	A	O	P	A	O	P	A	O
<i>z</i>		6.32	3.48	4.11	1.58	0.32	0.00	6.32	6.01	4.43
<i>P</i>		< 0.01	< 0.01	< 0.01	0.11	0.74	1.00	< 0.01	< 0.01	< 0.01

Table 4. *Kolmogorov-Smirnov two-sample tests of significance*

Compari- sons		Visual angle relationships					
		2:1		1:1		1:2	
		<i>D</i>	<i>P</i>	<i>D</i>	<i>P</i>	<i>D</i>	<i>P</i>
P-A		31	< 0.01	5	N.S.	23	< 0.01
P-O		29	< 0.01	5	N.S.	32	< 0.01
A-O		6	N.S.	5	N.S.	13	< 0.05

These trends held for all three sets of instructions: thus stimulus size and distance were variables even when the subject was instructed to use 'objective' judgemental criteria. It can be asked, however, whether variation in the sizes and distances of the stimuli has had more effect on the brightness judgements of those subjects who were instructed to use one set of criteria than on the judgements of subjects instructed to use other criteria. For each of the three visual angle relationships the cumulative frequencies of 'brighter' judgements with increase in the luminous intensity of the comparison stimulus can be compared for the three instruction groups. The comparisons have been made with the Kolmogorov-Smirnov two-sample test of significance (Siegel, 1956) and they involve comparing the 'projective' distribution with the 'apparent' distribution, the 'projective' with the 'objective', and the 'apparent' with the 'objective'. *D* values, and the probability of obtaining these by chance alone, are set out in Table 4. One-tail probabilities are given.

When the two stimuli subtended the same visual angle, brightness judgements were independent of variation in stimulus size and distance irrespective of the instructions given to the subject. When the stimuli did not subtend the same visual angle, the brightness judgements of subjects instructed to use 'projective' criteria were significantly more affected by variation in the sizes and distances of the stimuli



than were the judgements of subjects instructed to use either 'apparent' or 'objective' judgemental criteria. The differences in cumulative frequency distributions between the 'apparent' and the 'objective' instruction groups were not significant.

#### IV. DISCUSSION

The conclusions to be drawn from the present experiment are limited to the luminance values and the visual-angle relationships used; with certain experimental arrangements Hanes (1951) has found that a stimulus subtending a larger visual angle had to be of a greater luminous intensity than a stimulus subtending a smaller visual angle before the two stimuli were judged to be equal in brightness. For the values used in the present experiment it was found that brightness judgements given under reduced viewing conditions were some function of the sizes and distances of the stimuli, such that stimuli subtending smaller visual angles were judged to be less bright than stimuli subtending larger visual angles when the two stimuli were of the same luminous intensity. Even subjects who were explicitly instructed to make judgements, using criteria independent of variation in the sizes and distances of the stimuli, gave judgements which were some function of stimulus size and distance variation. Therefore, it cannot be said that results of the type obtained by Hanes (1951) are mere artifacts which have arisen because the subject explicitly used brightness equality criteria which were dependent upon stimulus size and distance.

Under reduced viewing conditions, and with subjects given 'objective' instructions, variation in stimulus size and distance has been shown to have the same molar effect on brightness judgements as variation in the luminous intensities of the stimuli has on size and distance judgements (Robinson, 1954; Coules, 1955). There are insufficient data to set out molecular relationships between the two sets of results.

It can be asked whether brightness judgements obtained under unrestricted viewing conditions are dependent upon stimulus size and distance. Size judgements obtained under unrestricted viewing conditions are independent of variation in stimulus distance (size-constancy); it is probable that, for subjects given 'objective' judgemental criteria, size judgements are also independent of variation in the luminous intensities of the stimuli under these viewing conditions. Distance judgements obtained under unrestricted viewing conditions are independent of variation in stimulus size (distance-constancy); it is probable that, for subjects given 'objective' judgemental criteria, distance judgements are also independent of variation in the luminous intensities of the stimuli under these viewing conditions. If the above relationships hold, and if the effect of variation in stimulus size and distance on brightness judgements parallels the effect of variation in the luminous intensities of the stimuli on size and distance judgements, it would be expected that brightness judgements given under unrestricted viewing conditions by subjects using 'objective' criteria would be independent of variation in the sizes and distances of the stimuli. If found, such a relationship would, by analogy with size-constancy and distance-constancy, be called brightness-constancy. As this term is commonly used to refer to a somewhat different relationship (Bartley, 1958), it is proposed to call the hypothesized relationship, whereby brightness judgements under unrestricted viewing

conditions are independent of variation in stimulus size and distance, luminance-constancy.

It was predicted that, for subjects instructed to use 'projective' criteria, as many judgements of 'brighter' as of 'less bright' would be given when the two stimuli subtended equal total amounts of light. The two stimuli subtended equal total amounts of light when, for the arrangements where the comparison stimulus subtended a visual angle twice as large as that subtended by the standard stimulus, the luminous intensity of the comparison stimulus was  $2 \cdot 20 \log \text{ ft.-L.}$ , and when, for the arrangements where the visual angle subtended by the comparison stimulus was half the visual angle subtended by the standard stimulus, the luminous intensity of the comparison stimulus was  $1 \cdot 40 \log \text{ ft.-L.}$  Two binomial tests of significance (both  $z = 3 \cdot 48$ ,  $p < 0 \cdot 01$ ) show that although brightness judgements given to 'projective' instructions are some function of stimulus size and distance, they do not approximate the above relationship. Subjects frequently commented on the difficulty of making judgements with these criteria and it was found after the experiment that few subjects were aware of the exact mathematical relationship between total light subtended and stimulus size and distance.

The finding that similar functions were obtained for subjects instructed to use 'apparent' and 'objective' criteria is of interest because of the widespread use of instructions in which subjects are told to judge in terms of 'apparent' brightness. Wright (1946, p. 76), for example, in describing the direct comparison method, states, 'The task of the observer then consisted in adjusting the brightness of the  $\lambda_2$  field until it appeared equal in brightness to that of the  $\lambda_1$  field under the given conditions of observation'. Here the subject may have ignored the difference in wavelength in making judgements, or he may have confounded wavelengths and luminous intensities in making judgements. If subjects are prone to confound luminous intensity and other variables in making brightness judgements with 'apparent' judgemental criteria, it would be probable that many of the basic photometric functions were artifacts; that is, that different functions would have been obtained if the subject had been given 'objective' criteria. The results of the present experiment show, however, that 'apparent' criteria are not ambiguous; that is, that they lead to no more confounding of luminous intensity with other variables than do 'objective' criteria. More studies are needed in which brightness discrimination functions are obtained with both 'apparent' and 'objective' judgemental criteria.

#### REFERENCES

- BARTLEY, S. H. (1958). *Principles of Perception*. New York: Harper and Brothers.
- COULES, J. (1955). Effect of photometric brightness on judgements of distance. *J. Exp. Psychol.* **50**, 19-25.
- GIBSON, J. J. (1950). *The Perception of the Visual World*. Boston: Houghton Mifflin.
- GILINSKY, A. S. (1955). The effect of attitude upon the perception of size and distance. *Amer. J. Psychol.* **68**, 173-92.
- HANES, R. S. (1951). Suprathreshold area brightness relationships. *J. Opt. Soc. Amer.* **41**, 28-31.
- ITTELSON, W. H. & KILPATRICK, F. P. (1961). The monocular and binocular distorted rooms. In Kilpatrick, F. P. (ed.), *Explorations in Transactional Psychology*. New York University Press.
- OVER, R. (1960). Size and distance judgements under reduced viewing conditions. *Aust. J. Psychol.* **12**, 162-8.

- OVER, R. (1960*a*). The effect of instructions on size-judgments under reduced viewing conditions. *Amer. J. Psychol.* **73**, 599-602.
- OVER, R. (1961). Distance-constancy. *Amer. J. Psychol.* **74**, 309-10.
- OVER, R. (1962). Stimulus wavelength variation and size and distance judgements. *Brit. J. Psychol.* (in the Press).
- ROBINSON, E. J. (1954). The influence of photometric brightness on judgements of size. *Amer. J. Psychol.* **67**, 464-74.
- SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw Hill.
- THOULESS, R. H. (1931). Phenomenal regression to the real object. I. *Brit. J. Psychol.* **21**, 339-59.
- WRIGHT, W. D. (1946). *Researches on Normal and Defective Colour Vision*. London: Henry Kimpton.

(*Manuscript received 10 October 1961*)



## THE PERCEPTION OF RHYTHMICALLY REPEATED LINEAR MOTION IN THE HORIZONTAL PLANE

By E. G. WALSH

*Department of Physiology, University of Edinburgh*

Normal subjects were subjected to linear motion rhythmically repeated at rates in the range 1 to  $\frac{1}{4}$  eye. sec. With horizontal oscillations of  $\frac{1}{2}$  eye. sec. the sensations of moving came at the turning-points of the swing. The person felt he was moving most rapidly when he was momentarily stationary. The C.N.S. interprets acceleration as velocity under these circumstances. At times the first sensation of moving in one direction came whilst the person was travelling rapidly in the opposite way. At 1 eye. sec. the sensations were in time with the movements as they occurred. Threshold data indicate that the peak acceleration adequately describes the stimulus at  $\frac{1}{4}$  and  $\frac{1}{2}$  eye. sec. The findings are discussed in relation to the behaviour of the otolith organs and may have a bearing on possible explanations of motion sickness.

### I. INTRODUCTION

The sensitivity of the human being to rhythmically repeated linear motion has been studied very little since Mach's work (Mach, 1875). This is surprising, for rhythmically repeated linear motion is produced both by normal locomotion and by many forms of transport. The principal purpose of the present investigation has been to obtain information about the sensitivity of the human being to oscillatory motion of different periodicities and to investigate the phase relationship between stimulus and sensation. Apparatus has been developed to provide motion both silently and with freedom from vibration, thus avoiding confusing sensory cues. A preliminary account of this work has already been published (Walsh, 1961*a*).

In so far as the sensibility to the motion is dependent upon the otolith organs, it is generally accepted that a displacement of the calcareous particles is involved. This distortion must be due to acceleration and it is sometimes assumed that the peak acceleration reached can be regarded as an effective index of the stimulus. This is likely to be true only under certain restricted circumstances, for the duration during which the acceleration acts and the rate at which it is applied and withdrawn must also be considered. The peak acceleration involved in rhythmic motion may be expected to be a guide to the efficacy of the stimulus only if the oscillations are both slow as compared with any natural resonance displayed by the mechanical structures of the end organ and are rapid compared with the rate of accommodation of the sensory endings. It appeared possible therefore that a frequency band delimited at its upper and lower ends by these factors might be found in which the strength of a stimulus could be adequately described in terms of the peak acceleration reached at the turning-points of the swing.

### II. METHODS

Initial observations made use of a trolley pulled backwards and forwards on rails by a motor-driven crank. It soon became clear that very small irregularities in the movements of the wheels on the rails could be disturbing to the subject. Accordingly, different systems of suspension were adopted and devices were built that would oscillate silently on their own for a number of cycles after an initial displacement without the application of further external force.

Movements of 1 to  $\frac{1}{2}$  eye./sec. were produced by a parallel swing (Jongkees & Groen, 1946). This consisted of a stretcher suspended at its four corners by wires; by varying the length of the wires the periodicity could be altered. For slower oscillations the system was modified, beams were attached by universal joints to the stretcher at one end and to the wall at the other (Fig. 1) after the manner of the ballistocardiograph described by Henderson (1905). By varying the points of attachment of the wires to the ceiling different rates of oscillation, down to  $\frac{1}{2}$  eye. sec., could be obtained. With wires 6 m. long movements of up to 2 m. amplitude could be obtained in the space available.

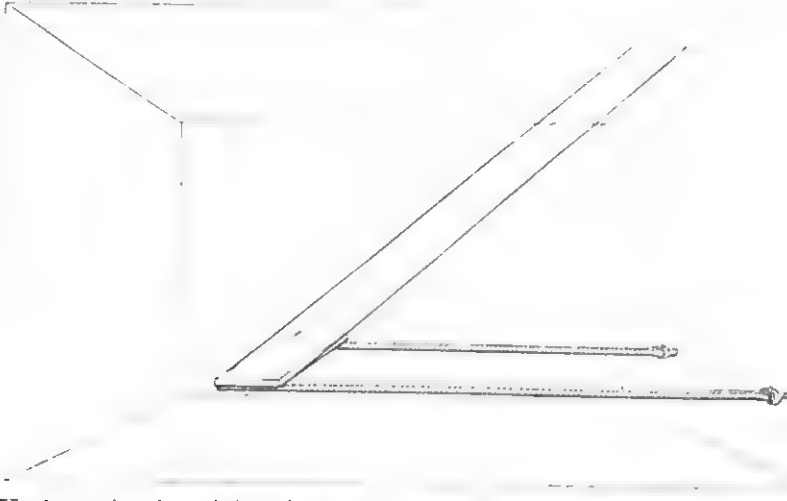


Fig. 1. Horizontal swing giving slow oscillations. Universal joints are attached to each end of the beams, only those at the ends fixed to the wall are depicted.

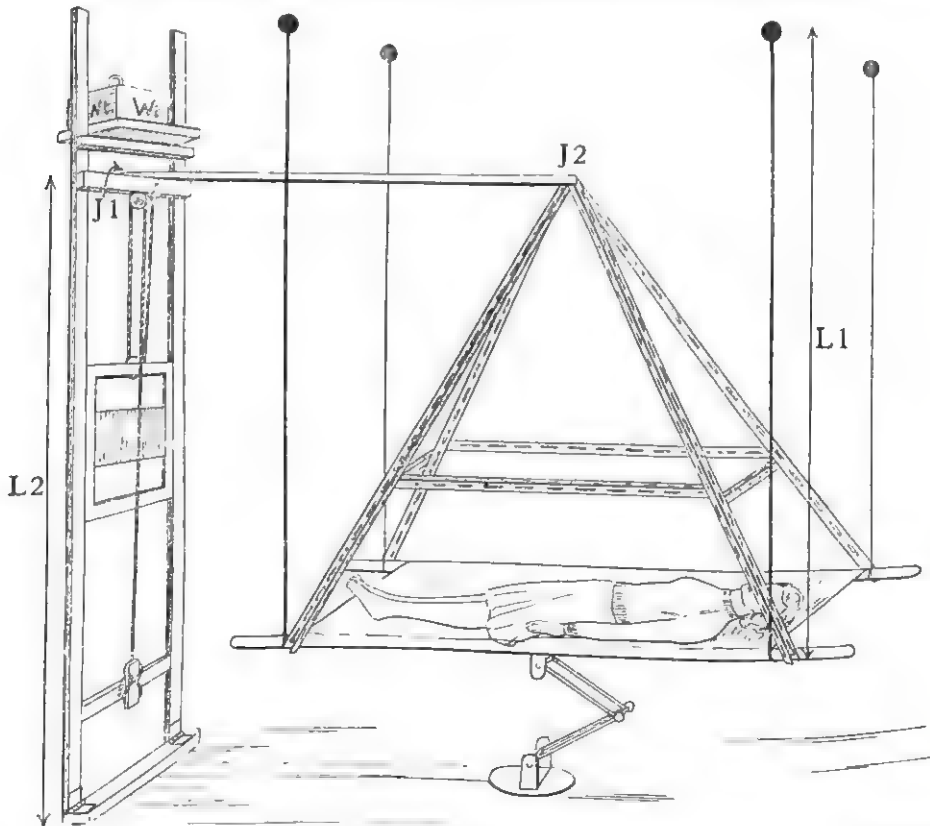


Fig. 2. Stretcher supported by four wires is linked to counterpoise by a steel bar. The counterpoise is hinged at its base. The frame attached to the underside of the stretcher restrains the motion of the stretcher to one plane.

In another device the parallel swing was again used; its motion was slowed by an inverted pendulum (Fig. 2). The pendulum was hinged to the floor and carried a heavy weight (Wt.) at its upper end. The pendulum was connected by a horizontal steel bar to a frame attached to the stretcher. The bar was furnished with joints (J 1 and J 2) at its ends. The distance from the hinges on the floor to the joint attaching the bar to the pendulum (J 1) equalled the length of the wires ( $L_1 = L_2$ ). For adjustment purposes the pendulum was furnished with a second weight which could be raised and lowered by a rope. Both types of apparatus for producing slow oscillations were used for the observations on phase relationships; the Henderson device alone was used for the threshold observations.

The sinusoidal nature of the movements was checked graphically by observing the records obtained from a kymograph equipped with a lever to record the motion. Values of acceleration quoted have been obtained by calculation according to the frequency and amplitude of the motion. The movements of the various devices involved no rotatory components. The motion was curvilinear but the departure from linearity was slight and in the calculations the corresponding radial components of acceleration have been neglected.

The subjects were young adult laboratory workers or medical students, apart from the children whose data are included in Table 1. The person lay on a plastic foam mattress, precautions being taken to prevent him feeling the movement of the air relative to his body. Thus, where appropriate, he lay in a sleeping bag totally enclosed except for his face. This was covered by a cloth, which also served as a blindfold. In all of the experiments the motion was in the long axis of the body.

### III. RESULTS

#### (1) Thresholds

The stretcher was started oscillating with an amplitude sufficiently great as to be readily detectable. With large enough movements the person always became aware of sensations arising in his skin and deeper tissues. Force is transmitted to the mass of the body through the surface during the oscillations, so this was expected. As the motion was allowed to die away this effect became less and less prominent. Eventually the person was usually left with a sense of movement with no associated somatic sensations. The least motion that could reliably be detected in this way has been taken as the sensory threshold.

Table 1. *Thresholds for the perception of rhythmically repeated horizontal motion*

Subjects—three young adults and two 10-year-old children. With each subject the mean of four determinations was taken as 'threshold'. The values in this table have been averaged over the different subjects, the standard deviations refer to the 'between subjects' variance.

Frequency (cyc./sec.)	Position				
	Supine	Prone	On right	On left	
$\frac{1}{2}$	3.7 $\pm$ 2.7	3.2 $\pm$ 3.1	3.0 $\pm$ 2.8	3.0 $\pm$ 2.7	Threshold expressed as amplitude of motion (cm.)
$\frac{1}{3}$	50.2 $\pm$ 28.5	25.2 $\pm$ 15.4	42.5 $\pm$ 19.4	31.8 $\pm$ 22.9	
$\frac{1}{4}$	3.9 $\pm$ 2.8	3.4 $\pm$ 3.2	3.1 $\pm$ 2.9	3.1 $\pm$ 2.8	Threshold expressed as peak velocity (cm. sec.)
$\frac{1}{5}$	17.5 $\pm$ 9.9	8.8 $\pm$ 5.4	14.8 $\pm$ 6.8	11.1 $\pm$ 8.0	
$\frac{1}{6}$	8.2 $\pm$ 5.9	7.0 $\pm$ 6.8	6.6 $\pm$ 6.1	6.6 $\pm$ 5.9	Threshold expressed as peak acceleration (cm. sec. <sup>2</sup> )
$\frac{1}{8}$	12.2 $\pm$ 6.9	6.1 $\pm$ 3.7	10.3 $\pm$ 4.7	7.7 $\pm$ 5.6	
$\frac{1}{10}$	17.1 $\pm$ 12.4	14.7 $\pm$ 14.2	13.7 $\pm$ 12.8	13.7 $\pm$ 12.4	Threshold expressed as peak values for rate of change of acceleration (cm. sec. <sup>3</sup> )
$\frac{1}{12}$	8.5 $\pm$ 4.8	4.3 $\pm$ 2.6	7.2 $\pm$ 3.3	5.4 $\pm$ 3.9	

Reliable determinations of threshold could not be obtained at the faster rates of oscillation because of the very small distances involved and because of the disturbing effects of the ballistocardiogram. Technically satisfactory results were, however,



obtained with periods of 3 and 9 sec. The results with five subjects have been used as the basis for Table 1. In comparing the values obtained at  $\frac{1}{3}$  and  $\frac{1}{9}$  cyc./sec., substantial differences were found when the values were expressed in terms of amplitude, velocity, or 'jolt' (rate of change of acceleration). The values were more nearly similar when expressed in terms of peak acceleration. To obtain further evidence one subject, not included in the previous series, was studied in more detail. Four observations were made under each set of circumstances and the average value obtained. The procedure was repeated on six separate occasions and the data are summarized in the top half of Table 2. It is evident that there is a close correspondence between the threshold values at the different periodicities when the comparison is in terms of peak acceleration. The whole procedure was later repeated and the results are summarized in the bottom half of Table 2. Again the thresholds at the two frequencies are similar when expressed in terms of peak acceleration. The values as a whole became lower in this series, evidently an expression of the effects of training.

Table 2. *Thresholds for the perception of horizontal movement*

On each occasion the mean of four determinations was taken as 'threshold'. The procedure was repeated on six different days and the average values are shown. The standard deviations refer to the 'between days' variance. (Subject E.G.W.)

Frequency (cyc./sec.)	Position				
	Supine	Prone	On right	On left	
First set of results					
$\frac{1}{3}$	3.3 $\pm$ 1.4	4.9 $\pm$ 1.6	3.7 $\pm$ 0.6	3.7 $\pm$ 1.1	Threshold expressed as amplitude of oscillation (cm.)
$\frac{1}{9}$	32.8 $\pm$ 7.0	43.2 $\pm$ 18.9	35.0 $\pm$ 10.3	39.4 $\pm$ 11.9	
$\frac{1}{3}$	7.3 $\pm$ 3.0	10.8 $\pm$ 3.6	8.1 $\pm$ 1.3	8.1 $\pm$ 2.5	Threshold expressed as peak acceleration (cm./sec. <sup>2</sup> )
$\frac{1}{9}$	8.0 $\pm$ 1.7	10.5 $\pm$ 4.6	8.5 $\pm$ 2.5	9.6 $\pm$ 2.9	
Second set of results					
$\frac{1}{3}$	1.3 $\pm$ 0.3	1.7 $\pm$ 0.3	1.6 $\pm$ 0.3	1.5 $\pm$ 0.3	Threshold expressed as amplitude of oscillation (cm.)
$\frac{1}{9}$	14.8 $\pm$ 4.0	16.0 $\pm$ 2.3	15.7 $\pm$ 3.0	15.2 $\pm$ 3.9	
$\frac{1}{3}$	2.9 $\pm$ 0.7	3.7 $\pm$ 0.7	3.5 $\pm$ 0.7	3.3 $\pm$ 0.7	Threshold expressed as peak acceleration (cm./sec. <sup>2</sup> )
$\frac{1}{9}$	3.6 $\pm$ 1.0	3.8 $\pm$ 0.6	3.8 $\pm$ 0.7	3.6 $\pm$ 0.9	

## (2) *Phase relationships*

The phase relationships were investigated with motion slightly above threshold in eight subjects. With movements of lesser amplitude the person was sometimes aware that he was moving but could not reliably determine the direction. With oscillations of 1 cyc./sec. the person was required to indicate with a finger the direction in which he felt himself moving. The indications were accurate; thus the finger would point in the sense appropriate for the motion that was occurring at that time. With slower motions this was no longer true, and the movements of the finger tended to anticipate the motion of the stretcher. Thus with an oscillation of  $\frac{1}{3}$  cyc./sec. the finger movements were usually 90° ahead of the corresponding movements of the person's body. When the person indicated verbally the point at which he felt he was travelling most rapidly to the head, or to the feet, he was most commonly stationary at the turning-points of the swing; the movements to the head or to the feet had not then started. Similar results were obtained with slower movements, the phase shift between  $\frac{1}{3}$  and  $\frac{1}{9}$  cyc./sec. remaining at 90°. In absolute terms the anticipatory nature of the sensa-

tions varied from 0.75 sec. at  $\frac{1}{3}$  cyc./sec. to 2.25 sec. at  $\frac{1}{9}$  cyc./sec. Naturally the effect was much more striking to observe at the slower rates of oscillation. No difference was found when the person lay prone, or on one side rather than supine.

In other observations the person was asked to call out as soon as he felt motion starting in a particular sense. With a period of 9 sec. the initial sensation frequently came whilst he was still travelling rapidly in the opposite direction. In some observations the subject wore no blindfold and was instructed to open his eyes as soon as he had responded. He was liable then to be surprised to find himself travelling quickly in the direction opposite to the one his senses had indicated.

### (3) *Reaction times*

The interpretation given to the timing of the sensations is that of a phase advance with respect to the movement itself. At first sight it might be supposed that an equally valid interpretation would be to suppose that there existed a phase retardation of  $\frac{3}{4}$  cycle. With an oscillation of  $\frac{1}{9}$  cyc./sec. this would mean that the sensation was set up after a very long latent period from the time of maximum application of acceleration. No evidence for such long after-effects could be found. If the stretcher was stopped oscillating sensations were aroused by the deceleration at once and they were, moreover, evanescent, appearing to last no more than a fraction of a second. If, whilst the stretcher was oscillating, its motion was artificially increased or decreased, the person was usually at once aware of the interference. It was evident that the time elapsing between the alteration of velocity and the sensation was brief. To obtain numerical data on this point measurements of reaction time were made.

The subject lay supine upon the stretcher of a parallel swing. This was displaced in a feetward direction from its position of rest, held there for a variable interval and then released. The person was asked to press a morse key as soon as he felt the motion. Records of the movement and of the responses were obtained on a kymograph. The natural oscillations of the swing had a period of 3 sec. but the response was complete before the movement had swung far in a headwards direction. Two subjects were used. The first subject tested with a displacement corresponding to 18 cm./sec.<sup>2</sup> gave an average reaction time over four trials of  $0.52 \pm 0.09$  sec. The second subject tested under similar conditions for eight trials gave an average reaction time of  $0.31 \pm 0.10$  sec. When the peak acceleration was reduced to 9 cm./sec.<sup>2</sup> there was no elongation of the interval, the mean value over eight trials being  $0.29 \pm 0.08$  sec.

### (4) *Lesions of the vestibular apparatus*

It has been shown that after unilateral labyrinthectomy the threshold for abruptly applied horizontal acceleration varies with the orientation of the head (Walsh, 1960). The threshold is elevated when lying on the damaged side. If such an effect could be observed with slow oscillations some evidence would be afforded that the perceptions of the movement were dependent upon vestibular rather than somatic receptors. Miss 'B', a subject with a unilateral lesion of the vestibular system, discussed in the previous work, was tested on the 9 sec. swing using horizontal motion. The mean and standard deviation of twelve determinations when lying on the side of the lesion were  $18.5 \pm 3.0$  cm./sec.<sup>2</sup>. The values obtained from the same number of observations when

lying on the sound side were  $14.2 \pm 3.0$  cm./sec.<sup>2</sup>. This difference is highly significant ( $t = 3.51$ ;  $P < 0.005$ ). Two other subjects with unilateral lesions were tested using horizontal motion and their thresholds too were elevated when lying on the side of the lesion. For horizontal oscillations with a period of  $2\frac{1}{2}$  sec. the data already presented (Walsh, 1961*b*) indicate the primacy of the role of the vestibular apparatus for the perception of that motion.

#### IV. DISCUSSION

Both for horizontal and vertical movements the sensations must be regarded as biologically most satisfactory with the fast (1 cyc./sec.) motion, for under these conditions the person feels he is travelling most rapidly in a certain direction when that movement is in fact taking place. With slower periodicities, however, the person feels he is moving most rapidly in a given direction when he is stationary at the turning-point of the oscillation. The first feeling that motion is taking place in one direction frequently arises when he is in fact travelling rapidly in the opposite direction. The central delay cannot be estimated but is likely to be rather shorter than the reaction time found in the present experiments to be 0.29 to 0.52 sec. The supposition is that the C.N.S. interprets acceleration as velocity so offsetting neural delay. In using acceleration or its derivative as a guide to future events the C.N.S. is making use of a system familiar to those interested in servo systems. It is supposed that the phase advance is normally used to compensate for the time taken in central neural events. With the relatively rapid periodicities that occur in walking and running such a system may prove fairly satisfactory. When, however, with the aid of man-made devices the person is subject to relatively slow oscillations, the system has too much phase advance. Perhaps similar effects play a role in the disturbances of gait seen when people unaccustomed to the sea first try to walk on the deck of a rolling ship. Preliminary observations using motion in the vertical direction have shown that under these conditions too the sensations may be out of step with the movement. The periodicities of the motion which give rise to the misleading timing of the sensations are those to which motion sickness is attributed (for review see Tyler & Bard, 1949). When the eyes are open slow linear motion will give rise to a conflict of sensory cues.

In this discussion it has been assumed that the natural period of the otolith membrane is high compared with the periodicity of the most rapid motion used in the study. The properties would thus be similar to the fish otolith organs, the mechanics of which have been elegantly discussed by de Vries (1956). If it should later be shown that the natural period is lower than has been assumed the phase shifts shown by a resonant system when excited close to its natural period would also require consideration.

#### V. SUMMARY

1. Apparatus has been developed that allows a subject to oscillate in the horizontal plane without rotation. Movements with periods of 1-9 sec. have been studied. The person indicated the direction in which he felt the motion was occurring.
2. The timing of the sensations of movement appeared to be correct with oscillations of 1 cyc./sec.
3. For oscillations of  $\frac{1}{3}$  cyc./sec. or slower, the sensations of moving were in anticipation of the motion, there being a 'phase advance'. The first sensations of motion



in one direction were frequently aroused when the person was still travelling rapidly in the opposite direction.

4. Measurements of thresholds with cycles of 3 and 9 sec. duration show similar values for peak acceleration.

The author is indebted to the subjects for their patient co-operation. The apparatus was constructed by Mr G. Wright. The figures were prepared in the Department of Medical Illustration.

#### REFERENCES

- HENDERSON, Y. (1905). The mass movements of the circulation as shown by a recoil curve. *Amer. J. Physiol.* **14**, 287-98.
- JONGKEES, L. B. W. & GROEN, J. J. (1946). The nature of the vestibular stimulus. *J. Laryng.* **61**, 529-41.
- MACH, E. (1875). *Grundlinien der Lehre von den Bewegungsempfindungen*. Leipzig: Engelmann.
- TYLER, D. P. & BARD, P. (1949). Motion sickness. *Physiol. Rev.* **29**, 311-369.
- VRIES, H. L. DE (1956). Physical aspects of the sense organs. *Progr. Biophys.* **6**, 208-64.
- WALSH, E. G. (1960). Perception of linear acceleration following unilateral labyrinthectomy: variation of threshold according to orientation of the head. *J. Physiol.* **153**, 350-7.
- WALSH, E. G. (1961*a*). Sensations aroused by rhythmically repeated linear motion-phase relationships. *J. Physiol.* **155**, 53*P*.
- WALSH, E. G. (1961*b*). Role of the vestibular apparatus in the perception of motion on a parallel swing. *J. Physiol.* **155**, 506-13.

(Manuscript received 13 September 1961)



## PERCEPTUAL JUDGEMENT AS A FUNCTION OF CONTEXT

By LUDWIG IMMERGLUCK

*San Francisco State College*

Perceptual contrast has been traditionally demonstrated with stimulus attributes involving size, perspective, or brightness of given stimulus objects. The present investigation aims to examine the operation of perceptual contrast in such qualitative dimensions as geometric stability, symmetry, or general design quality of a stimulus.

The results show that the stimulus figure, a square, was perceived as geometrically more perfect when presented in a disordered and unstable context as compared with an ordered and stable one.

### INTRODUCTION

When a strip of neutral grey is held against a white background it is perceived as almost black, whereas it appears to be very light grey when a black background is substituted. The size of a circle is underestimated when presented in conjunction with a larger circle and overestimated when placed in contextual proximity to a smaller one. Perhaps no other aspect of perception illustrates the importance of context more emphatically than do the phenomena of perceptual contrast. Most of the familiar visual illustrations (Beeler & Branley, 1951) as well as a considerable array of experimental investigations (Ames, 1951; Ittelson & Kilpatrick, 1953; Lawrence, 1949; Witkin, Lewis, Hertzman, Machover, Meissner & Wapner, 1954) serve to demonstrate the role of context not only in shaping but also in distorting the perception of stimulus objects. Most of these studies, however, utilize contrast only incidentally or else as a means by which broader processes of perceptual distortion are illustrated. We are still left in the dark with regard to the nature of perceptual contrast as such. What specific processes underlie it? What attributes of the perceptual response come into play when a marked disparity between stimulus object and context exists? What are the conditions under which perceptual contrast will not take place? And, more generally, what role, if any, do these contrast phenomena play in everyday life perception? This latter formulation raises the question regarding the extent to which 'contrast distortions' affect more complex stimulus dimensions. Most contrast illustrations and pertinent experimental designs involve perceptual judgements of length, height, overall size or brightness of a stimulus object. What about such stimulus dimensions as symmetry, stability, or geometric order? Will, for example, the same stimulus object be perceived as more stable and ordered in a disordered as compared with a stable and ordered context? It is to this specific question that the study at hand addresses itself.

The present study constitutes a part of a series of investigations aimed at identifying specific attributes of perceptual responses. In the course of these investigations the writer has proposed a number of basic principles which are assumed to inhere in the perceptual process. One of these principles has been described as perceptual homeostasis. Briefly defined, it refers to a tendency of the perceptual response to establish equilibrium, to balance discrete parts of the perceptual field. This balancing process may be achieved through various means. At times a tendency is observed to render the diverse areas of a stimulus field into a harmonious or congruent total



percept, at other times the dissimilarities or contrasts between the different parts of the field (hence, between a stimulus and its context) become emphasized. Contrast phenomena are assumed to represent, then, one way in which the postulated homeostatic tendency of perception expresses itself. The implication here is that stimulus dimensions in one part of the perceptual field are counterbalanced by opposing or contrasting dimensions in another part of the field. Furthermore, it is assumed that this aspect of perceptual homeostasis underlies broad and more complex perceptual processes and should, therefore, be demonstrable in stimulus dimensions other than those involved in the judgement of length, size or brightness of a stimulus object.

Within the framework of the above considerations, the present investigation aims to demonstrate perceptual contrast in a discrimination task involving more complex stimulus dimensions. Specifically, such attributes as geometric stability and symmetry as they are affected by a contrasting context were put to an experimental test. In accordance with the perceptual hypothesis under discussion, the following results should be obtained: when subjects are asked to evaluate a stimulus figure from the point of view of geometric orderliness or stability, the same stimulus design should be perceived as more stable in a disordered and unstable as compared to a stable or orderly context.

## METHOD

### *Material*

The test material consisted of two cards  $4 \times 6$  in., a Stable Context Card and an Unstable Context Card. Each contained as the identical stimulus figure a square  $2 \times 2$  in. and a different context design. In the Stable Context Card the design consisted of three symmetrically arranged circles (radius 1.5 cm.) and in the Unstable Context Card of a cluster of asymmetrical and unrhythmical lines. In order to eliminate positional, area or directional differences between the context designs, they were placed in similar positions above the squares and covered approximately the same area (see Fig. 1).

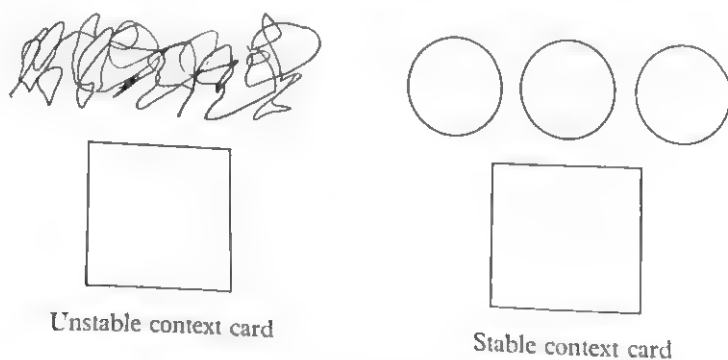


Fig. 1

### *Subjects and procedure*

The subject group was comprised of forty male undergraduate students ranging in age from 19 years through 25 years (male subjects were chosen because of some evidence (Witkin, 1950) for the existence of sex differences in perceptual organization). They were obtained from introductory classes in psychology and had no knowledge regarding the purpose of the experiment. None had any previous course work or

training in psychology and the experiment was conducted at the beginning of the course, when they had not yet been introduced to any class discussion or reading on perception.

Each subject was informed that he would be presented with two cards containing 'geometric designs'. He was told to pay attention only to the squares, that one of these squares would be slightly *better* than the other and that he was to identify the geometrically more perfect square. In the instructions it was emphasized to apply only geometric rather than aesthetic or other criteria in the judgement. The cards were then exposed side by side for 4 sec. and all judgements recorded. This exposure time was employed for two reasons: (1) In previous experiments (Immergluck, 1952) it was found that approximately this exposure time was needed for adequate perceptual appraisal of similar geometric designs; and (2) for the present study it was important to assess the subjects' initial perceptual organization which might have become altered through longer exposure time. In order to rule out any possible positional effects influencing the judgements the position of the cards was reversed for half the subject group.

# RESULTS AND DISCUSSION

The results summarizing the judgements for the total subject group are given in Table 1. As will be seen the difference in the judgements between the two contexts (measured by  $\chi^2$ ) is significant at the 0.1% level.

Table 1. *Perceptual judgements of the 'better square'*

	N	Unstable context	Stable context	Significance of difference
Total group	40	31	9	(Preference) $P < 0.001$
Unstable context, card on left	20	15	5	(Position) N.S.
Unstable context, card on right	20	16	4	

The results of the perceptual judgements as they were given when the respective position of the cards was reversed for half of the subject group are also presented in Table 1.

The difference between the two groups of judgements is statistically not significant and indicates that the spatial position of the cards had no bearing on the perceptual choice of the 'better square'.

The results indicate that the identical stimulus figure is perceived differently when placed in conjunction with a symmetrical context design as compared with a disordered and geometrically unstable design pattern. The square beneath the disorganized lines emerged perceptually as a 'better' and 'geometrically more perfect' figure. The findings appear, then, to be in accordance with the postulated perceptual principle in revealing the operation of a distortive process that favoured the figure in the unstable context by imbuing it with qualities of greater stability and sharper geometric organization.

The significant aspects of the present investigation centre around the particular nature of the stimulus dimensions involved. The results suggest that contrast phenomena and, more specifically, context-based perceptual distortions, not only are

operant in processes involving the perception of size or distance, but also participate in the perception of 'geometric' or 'design' attributes of stimulus objects. The extent to which contrast, and indeed context in general, influences other and perhaps even more complex stimulus attributes and the specific manner in which this influence is mediated constitute important questions that invite further research in this area.

#### SUMMARY

The study at hand was designed to examine the operation of contrast effects in the perceptual judgement of such stimulus attributes as 'geometric stability' or 'symmetrical quality' of simple stimulus designs.

Two squares of identical dimensions but in different contexts, one in an ordered and geometrically stable, the other in a disordered and structurally unstable one, were simultaneously exposed to subjects who were asked to identify the 'geometrically more perfect square'. In order to rule out the potential influence of positional effects, the spatial position of the two respective contexts was reversed for half of the subject group.

The square when presented in the disordered context was judged to be the geometrically more perfect and stable one. The results indicate that, at least within the confines of the present investigation, such stimulus attributes as symmetrical quality or geometric stability are subject to the same contrast phenomena or context-based distortions as are size, distance, or brightness dimensions of stimulus objects.

#### REFERENCES

- AMES, A., Jr. (1951). Visual perception and the rotating trapezoidal window. *Psychol. Monogr.* **65**, no. 324.
- BEELER, N. F. & BRANLEY, F. M. (1951). *Experiments in Optical Illusions*. New York: Crowell.
- IMMERGLUCK, L. (1952). The role of set in perceptual judgment. *J. Psychol.* **34**, 181-9.
- ITTELSON, W. H. & KILPATRICK, F. P. (1953). Experiments in perception. *Scientific American Reader*, pp. 576-96. New York: Simon and Schuster.
- LAWRENCE, M. (1949). *Studies in Human Behavior*. Princeton: Princeton University Press.
- WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. F. & WAPNER, S. (1954). *Personality Through Perception*. New York: Harper.
- WITKIN, H. A. (1950). Individual differences in ease of perception of embedded figures. *J. Pers.* **19**, 1-15.

(Manuscript received 5 September 1961)



## THE EFFECT OF RESERPINE ON CONDITIONED FEAR RESPONSES

By R. D. SAVAGE\*

*University of New England, Australia*

The effect of various amounts of the drug reserpine, a derivative of rauwolfia, were investigated in relation to foreleg flexion and ambulation responses in goats subjected to electric shock stimulation. Measures revealed that as the drug dosage increased the flexion scores increased and the ambulation decreased. In the control group, without the drug, the number of flexions and ambulation scores was maximal. The findings are interpreted in terms of the possible sites of action of the drug.

### I. INTRODUCTION

Investigators have reported the effects of the administration of rauwolfia derivatives on the behaviour of several species, including monkeys, dogs, cats, birds, fish, horses, cattle, mice, chinchillas and rats (Dews & Skinner, 1956; Earle, 1956; Uhr & Miller, 1960). Reserpine, a derivative of the root rauwolfia, is classified as a tranquillizer and produces psychological effects resembling those of the phenothiazines. Learned responses are affected by the tranquillizing properties of reserpine. Brady (1956) found that reserpine weakened conditioned emotional responses in rats and Weiskrantz (1957) obtained similar effects on conditioning in monkeys. Further results reported by Miller (1956) suggested that this drug modifies fear and avoidance responses, and Olds, Killam & Bach-y-Rita (1956) found that it suppressed behaviour directed by stimulation of the septal area. In the present experiment the tranquillizing effect of reserpine is measured in terms of the number of flexion responses to electric shock stimulation.

Animals receiving reserpine or chlorpromazine display a poverty of spontaneous movement. This is probably related to the stimulation of the limbic system. The hypothalamus is also stimulated by rauwolfia alkaloids and phenothiazines, whilst the reticular formation is slightly depressed by these drugs. It would seem that depression of the reticular formation and stimulation of the limbic system should reduce general activity. This possibility is measured in terms of the ambulation of animals with and without reserpine.

### II. SUBJECTS AND APPARATUS

The subjects were eight experimentally naïve young goats (average age 43 days) reared in identical conditions. The animals were housed with their mothers separate from the rest of the colony and allowed to eat and drink *ad libitum* for 30 min. each day, 2½ hr. before the experimental procedure began. They were weighed immediately after feeding each day, then given the appropriate injection.

A harness was attached to the goat and was connected to a dual pen recorder which traced a model of the animal's movement in the testing room. The movements of

\* The author is grateful to Ciba Laboratories, U.S.A., for supplying the drug 'serpasil' for this investigation.

the pens were calibrated to give a measurement in number of feet traversed. This figure was recorded on an automatic electronic counter. Electrodes were fastened to the right foreleg of the animal. Neither the harness nor the electrodes hindered the animals' movement about the testing room.

In this experiment the fear stimulation was an unconditioned stimulus of 10 V. a.c. of electric shock. A continuous 2 sec. electric shock was presented to the right foreleg of the animal.

### III. EXPERIMENTAL PROCEDURE

The animals were weighed and given the appropriate injection 2 hr. before the experimental session. A severe susceptibility to reserpine was found in the goats as compared with, for example, mice and rats, and it was only after several pilot studies that levels of drug dosage which allowed the animals to be used as experimental subjects were established. No report of the use of reserpine on goats is known to the author, but Earle (1956) has reviewed some of the effects of this drug in various species and points out differences in their tolerance of the drug.

Table 1. *Drug dosage and the mean flexion and ambulation scores*

Group	Drug dosage (mg./lb.)	Flexion scores	Ambulation scores
1	0.125 saline	20	320
2	0.0375 reserpine	12.75	260
3	0.0625 reserpine	16	181
4	0.125 reserpine	19.75	137

The doses of the drug used in this experiment are shown in Table 1. (It will be seen that Group 1 had no drug but instead 0.125 mg./lb. saline solution of 2.5 mg./c.c.) The drug was given in an oily solution of 2.5 mg./c.c.

Even with the doses used it was difficult to study the animals, especially those given the larger amounts of the drug. These goats were found particularly sensitive to the intramuscular injection of reserpine. Tremoring, soft stool and collapse to the ground with heavy breathing occurred with still higher doses, and were displayed after continued daily administration of the drug to the experimental animals.

In the testing situation the animals were given twenty trials per day for 2 days. Each trial consisted of 2 sec. of electric shock stimulation, with an interval of 58 sec. between successive presentations of the shock. The number of feet traversed throughout each trial (the 'ambulation score') was automatically recorded. The experimenter seated in the observation room outside the testing room noted the presence or absence of the conditioned foreleg flexion. The criterion of an unconditioned response was a subjectively judged 90° flexion of the right foreleg to the electric shock stimulation. The number of such flexions during a given trial constituted the 'flexion score'.

### IV. RESULTS

The animals were given four ambulation and four flexion scores, each the average of five trials. The raw data for the flexion responses and the ambulation scores showed homogeneity of variance and both sets of scores treated by analysis of variance.

The mean scores are shown in Table 1, and the respective analyses are set out in Tables 2 and 3. It will be seen that differences between groups for the flexion scores were significant at the 0.01 level. The higher the drug dosage the greater the number of flexions. The variances between trials were significant at the 0.05 level, but the interaction was not significant.

For the ambulation scores the differences between the groups were significant at the 0.05 level. The higher the drug dose the lower the ambulation score. The variances between trials and interaction between condition and trials were not significant.

Table 2. *Analysis of variance of foreleg flexion responses*

Source	S.S.	D.F.	M.S.	F	P
Variance between conditions	71	3	23.67	8.24	< 0.01
Variance between trials	29	3	9.67	3.37	< 0.05
Interaction between conditions $\times$ trials	52	9	5.78	2.01	N.S.
Residual	46	16	2.87		
Total	198	31			

Table 3. *Analysis of variance of ambulation scores*

Source	S.S.	D.F.	M.S.	F	P
Variance between conditions	106.476	3	35.472	3.73	< 0.05
Variance between trials	75.125	3	25.042	2.61	N.S.
Interaction between conditions $\times$ trials	100.536	9	11.170	1.17	N.S.
Residual	152.719	16	9.545		
Total	435.126	31			

## V. DISCUSSION

It can be seen from this experiment that the effect of reserpine on behaviour is to some extent dependent on the particular response required. The specific learning of a flexion response was increased, whilst general activity in terms of ambulation decreased as the drug dose received increased.

In this experiment, the number of flexion responses to an electric shock increased as the amount of drug increased, though the no drug group still showed the highest number of flexions. The learning of the flexion response between trials was significant within each group at the 0.05 level. The differences between groups were significant at the 0.01 level. However, Brady (1956), Miller (1956) and Weiskrantz (1957) have shown that conditioning is weakened by the administration of reserpine to rats and monkeys. Miller (1956) also has suggested a depressant effect of reserpine on fear and avoidance responses. These results probably related to the central depressant effect of the drug. The apparent contradiction between the present results and those of previous investigators may be a function of three main factors, the species of animal, the relative drug dosages used and the response required. For example, foreleg flexion to electric shock is probably a simple spinal reflex, not involving complex learning. Consequently, one might expect the lowered C.N.S. effectiveness to have a less detrimental effect than on the more complex learning in the Miller (1956), Brady (1956) and Weiskrantz (1957) experiments.



The ambulation scores were significantly lower the larger the drug dose of the group, but were not significantly different between trials within groups. This lowering of general activity could be related to the effect of the drug on the limbic, hypothalamic and reticular systems, or a general decrease in motor facility because of the drug's C.N.S. depressant effect. The problem is yet to be resolved.

The fact that general activity was lowered and specific response performance increased with the larger drug doses used in this investigation suggests the possible use of this drug for re-learning in experimental and clinical settings.

#### REFERENCES

- BRADY, J. V. (1956). The assessment of drug effects on emotional behaviour. *Science*, **123**, 1033-4.
- DEWS, P. B. & SKINNER, B. F. (1956). Techniques for the study of behavioral effects of drugs. *Ann. N.Y. Acad. Sci.* **65**, 247-356.
- EARLE, A. E. (1956). Reserpine (serpasil) in veterinary practice. *J. Amer. Vet. Med. Ass.* **129**, 227-33.
- MILLER, N. E. (1956). Effects of drugs on motivation: the value of a variety of measures. *Ann. N.Y. Acad. Sci.* **65**, 318-23.
- OLDS, J., KILLAM, K. F. & BACH-Y-RITA, P. (1956). Self-stimulation of the brain used as a screening method for tranquilizing drugs. *Science*, **124**, 265-6.
- UHR, L. & MILLER, J. G. (1960). *Drugs and Behavior*. New York: Wiley.
- WEISKRANTZ, L. (1957). Reserpine and behavioral non-reactivity. In *Psychotropic Drugs*, S. Garattini and V. Ghetti (eds.). New York: Elsevier.

(Manuscript received 11 October 1961)

## CORRESPONDENCE

*From Prof. H. J. Eysenck*

Foulds's recent attempt (this *Journal*, 1961, **52**, 385-7) to demonstrate 'the logical impossibility of using hysterics and dysthymics as criterion groups in the study of introversion and extraversion' is curiously reminiscent of the many philosophical attempts to demonstrate the logical fallacy of inductive scientific reasoning. Just as the scientist has to put up more or less cheerfully with the philosopher's contumely, and maintain his belief that his laws and theories do in fact have an empirical referent, so in this case also we may be able to feel that the picture is not quite as black as it is painted by Foulds. This feeling derives some justification from the fact that Foulds appears to misconceive the role of logic in this whole enterprise. What he appears to argue is that there is no logically watertight deduction from Jung's theory to the use of hysterics and dysthymics as criterion groups; as he puts it, 'the assumed correlation has to be demonstrated'. With this statement one cannot reasonably argue: Jung's hypothesis, and the deductions made from it, obviously require empirical support, and much work has gone into the attempt to furnish such support (see my *The Structure of Human Personality*). What is difficult to see is why Foulds should imagine that he was doing anything but stating the obvious, or that the present writer would hold a view different from his. The usual procedure has been to show that tests which according to theory should discriminate between extraverts and introverts (as diagnosed by the M.P.I. or some similar device) also discriminate between hysterics and dysthymics (as diagnosed by experienced psychiatrists). Franks's demonstration (*J. Abnorm. Soc. Psychol.* 1956, **52**, 143-5; *Brit. J. Psychol.* 1957, **48**, 119-26) that as extraverts are less easy to condition than introverts so hysterics are less easy to condition than dysthymics, is one of many examples of this method. The results certainly seem to verify the truth of what Foulds call 'Eysenck's equation', i.e. extraversion: introversion = hysteria/psychopathy: dysthymia. Even if the deduction were prey to logical errors, which cannot be regarded as demonstrated by Foulds, the empirical verification would nevertheless stand unimpaired as a factual contribution to science (see my *Experiments in Personality*, 1960; *Experiments with Drugs*, 1962). Having demonstrated that factual deductions can successfully be made from 'Eysenck's equation', we are entitled to apply it to the verification of new and different hypotheses, and to make use of its demonstrated value for other purposes, as in the use of hysterics and dysthymics as criterion groups for extraversion-introversion. To deny this is to deny the very existence of inductive reasoning.

Foulds also argues against a 'second deduction from Jung's hypothesis', to wit, that normal groups should be intermediate between hysterics and dysthymics with respect to extraversion-introversion. It is certainly possible to make up very complex and unlikely models, having no relationship to Jung's own theoretical system, to which Foulds's argument would formally apply. What I attempted, in my early writings, was to put Jung's rather unsystematic and descriptive writings into a form in which they could be tested experimentally; if Foulds is dissatisfied that this remodelled theory is sufficiently in line with Jung's own presentation, then he is at liberty to regard these statements simply as postulates in my theory, rather than as deductions from Jung. What is at issue, surely, is the simple question of whether or not randomly selected normal groups are or are not, as a matter of fact, intermediate between hysterics and dysthymics on the E-I dimension; the precise degree of the writer's indebtedness to Jung is of purely academic interest. Here also the facts seem to support the hypothesis with considerable force (Eysenck & Claridge, *J. Abnorm. Soc. Psychol.* 1962, to appear).

All this is not to say that quite genuine problems do not arise when one attempts to make use of well-known theories like Jung's or Pavlov's theory relating inhibition to hysteria. In planning one's research, one must pay respectful attention to what these great figures of the past have to say; yet the oracles speak at times mysteriously, inaudibly and even contradictorily. It is a salutary experience to go through a painstaking record of the many changes which such a theory as Pavlov's underwent in its application to psychiatry; almost any possible correlation between psychiatric syndrome and Pavlovian concept is postulated at one time or another.

However one may try to relate these pronouncements to one's own work, it is difficult to avoid the criticisms of either using a great name to justify one's own immature and incorrect theorizing, or else of putting forward as one's own theory something postulated many years ago by one of these intellectual giants. For the sake of historical accuracy, and as a fitting token of respect, I have always given credit to Jung and Pavlov for important parts of my system of postulates and deductions; I am not, however, very much concerned with the logical relations between the original theories, in one of their various forms, and my own, and I cannot feel that such concern would justify spending much time and energy in defending my usage against Foulds's criticisms.

Even as a philosophical *divertissement*, however, Foulds's argument would be more convincing if it were not marred by logical errors. Thus he writes of 'Eysenck's equation of dysthymia and introversion', when it must surely be clear that 'Eysenck's equation' is one of *relations*, not of *fundaments*. To say that:

$$\text{Black:White} = \text{High:Low}$$

is correct in that the relation of *oppositeness* is properly asserted to apply equally to both sides of the '=' sign; it is not asserted that black = high, or white = low! I do not *equate* dysthymia and introversion, but assert that the relation between introversion and extraversion is similar to that between dysthymia and hysteria, or, more technically, that the projection in *n*-dimensional space of the positions of dysthymic and hysteric groups on the E-I axis cuts this axis at points displaced from the population mean towards the I and the E ends, respectively. Clearly hysterics and dysthymics as groups are not located *on* this axis, as are extraverts and introverts, because they have high scores on neuroticism and are thus displaced along at least one major axis from normal E and I groups; this fact, which is a crucial part of my theory, and probably of Jung's also, would clearly prevent any straightforward equation of dysthymics and introverts. Indeed, I have explicitly made this clear by pointing out (in *Experiments in Personality*) that my hypothesis 'does not postulate a one-to-one correspondence... it preserves the distinction between personality, on the one hand, and symptomatology on the other'.

In summary, I would submit that the whole argument in terms of 'logical impossibilities' is misconceived; it is more appropriate to talk about 'the heuristic convenience of using hysterics and dysthymics as criterion groups in the study of introversion and extraversion'. We are concerned with facts and observable relationships, not with logic; a true appraisal of a theory can only come from the correspondence (or lack of correspondence!) of its postulates and deductions with observation and experiment.

### *From Dr G. A. Foulds*

Eysenck states, 'Jung's hypothesis, and the deductions from it, obviously require empirical support, and much work has gone into the attempt to furnish such support'. My point, which has not been met, was that no amount of work will clarify the issue within the framework of using hysterics and dysthymics as criterion groups for extraversion-introversion. This is a logical issue, not an empirical one.

Eysenck's 'usual method of procedure has been to show that tests which according to theory should discriminate between extraverts and introverts (as diagnosed by the M.P.I. or some similar device) also discriminate between hysterics and dysthymics (as diagnosed by experienced psychiatrists)' -notwithstanding that elsewhere (cf. the *Handbook of Abnormal Psychology*, p. xi) he states that psychiatric diagnoses are thoroughly unreliable. According to what theory? If the reference be to a theory other than the one I was questioning, it is irrelevant to my argument.

It was not being argued that there are no grounds other than the hysteric-dysthymic distinction for supposing that the M.P.I. E scale is measuring extraversion-introversion. There probably are such grounds and they are probably as adequate as we normally have in Psychology. What Eysenck could do is to argue the case for the construct validity of his scale and then demonstrate that hysterics as a group tend to be extraverted and dysthymics tend to be introverted. To show that an extreme degree of introversion or of extraversion made one more vulnerable to neurotic breakdown, it would be necessary to demonstrate that normals were normally distributed on E-I; whereas neurotics were bimodally distributed (or at least that they had a much larger standard deviation). Alternatively, normals could be classified as extraverts and introverts on the basis of some outside criterion, such as agreed observer ratings. It would then



be necessary to show that hysterics had higher scores on the M.P.I. E scale than did normal extraverts and that dysthymics had lower E scores than did normal introverts.

My objection, which has not been met, was that one cannot assume that the M.P.I. E scale measures extraversion-introversion merely because hysterics differ from dysthymics. If dysthymics had scored higher than hysterics, would the lower end of the scale have been designated 'extra-version'? Franks (1956, *J. Abnorm. Soc. Psychol.* 52, 143-50) concluded that anxious subjects were more easily conditioned in so far as they were more introverted. It would appear that conditionability is a function primarily of personality characteristics and secondarily of symptomatological characteristics; but instances of the interchangeability of introversion and dysthymia are numerous in the Eysenckian literature.

Eysenck says that: 'Black:White = High:Low is correct in that the relation of *oppositeness* is properly asserted to apply equally to both sides of the "=" sign; it is not asserted that black = high, or white = low! The writer is not *equating* dysthymia and introversion, but is asserting that the relation between introversion and extraversion is similar to that between dysthymia and hysteria.'

Eysenck claims that he is not equating hysteria and extraversion. On p. 10 of *The Handbook of Abnormal Psychology* he quotes one of my papers in support of the view that extraverts have high progressive matrices: Mill Hill Vocabulary ratio. What, in fact, I found was that hysterics had, relative to dysthymics, a high PM:MHV ratio when the PM was given with a 20 min. time limit and a low PM:MHV ratio when the PM was given without a time limit. I never mentioned extraverts or introverts.

His statement certainly does not say that black = high, or that extraversion = hysteria. It does not say anything whatsoever about the relationship between extraversion and hysteria. We cannot, therefore, predict that hysterics will come out on the E side and that dysthymics will come out on the I side, any more than we could predict that highs will come out on the black side and lows on the white side. In order to make such predictions Eysenck must say more than  $E:I = H:D$  in the sense in which he is now using this equation. What he in fact said, was: 'Jung maintains that much the most frequent neurotic disorder of the extraverted type is hysteria . . . the typical neurotic disorder of the introvert is psychasthenia (or dysthymia). Two consequences would seem to follow from this hypothesis. In the first place, among neurotics, those of a dysthymic pathology should be more introverted than those of a hysteric or psychopathic pathology . . . This is the crucial deduction on which hinges the usefulness of these nosological groups for the purpose of questionnaire validation. Formally, the argument may be put in the form of an equation:

$$\text{extraversion: introversion} = \text{hysteria/psychopathy:dysthymia.}$$

It is difficult to construe this to mean that extraversion bears the same relationship to introversion (presumably oppositeness) as hysteria bears to dysthymia *and nothing more*.

Eysenck has pointed out that his hypothesis 'does not postulate a one-to-one correspondence . . . it preserves the distinction between personality, on the one hand, and symptomatology on the other'. It preserves the distinction between personality and such symptomatology as is common to all neurotics; it confuses the distinction between personality and such symptomatology as is specific to neurotic subclasses, such as hysteria and dysthymia.

Penultimately, 'we are concerned with facts and observable relationships, not with logic'; but relationships are not observable. They have to be deduced - logically or illogically.

Ultimately it should perhaps be pointed out that, in *Psychological Types* (Kegan Paul, 1924, p. 288), Jung argues that, when an extravert actually develops hysteria, he becomes an introvert and, when an introvert actually develops psychasthenia, he becomes an extravert. For once Jungians and Freudians are agreed, since Fenichel and Glover both claim that conversion hysterics are introverts. Clearly then, if it were permissible to use hysterics and dysthymics as criterion groups for extraversion and introversion and if one were to follow Jung, hysterics should be the criterion group for introversion and dysthymics for extraversion. I am *not*, therefore, satisfied that 'this remodelled theory is sufficiently in line with Jung's own presentation' and I am grateful to Eysenck for leave 'to regard these statements simply as postulates in (his) theory'.

*From Dr J. G. Ingham, M.R.C. Social Psychiatry Research Unit, Llandough Hospital, Penarth*

Foulds contends that Eysenck makes an erroneous deduction from two hypotheses stated by Jung (1923). The hypotheses are that (1) 'much the most frequent neurotic disorder of the extraverted type is hysteria' and that (2) 'his (the introvert's) typical neurotic disorder is psychasthenia'. Eysenck identifies psychasthenia with his own term 'dysthymia' and the statement that Foulds considers an erroneous deduction is: 'among neurotics, those of a dysthymic pathology should be more introverted than those of a hysteric or psychopathic pathology'.

A source of confusion is that, while Jung's hypotheses are stated in terms of the qualitative types extravert and introvert, Eysenck introduces quantity by using the expression 'more introverted'. The 'more' might be taken to refer to the proportion of introverts. Thus interpreted, his statement becomes 'the proportion of introverts amongst dysthymics is higher than that amongst hysterics'. To see all the implications of the three statements it should be recognized that they refer to relationships within a  $2 \times 2$  contingency table. Such a table can be formed from the hypothetical figures given by Foulds (see Table 1) which, as he points out, are in accordance with Eysenck's statement but not with Jung's first hypothesis.

Let us consider the following general table of this type (which also includes Foulds's hypo-

	Extraverts	Introverts	
Hysterics	$a$ (160)	$b$ (40)	$a + b$ (200)
Dysthymics or psychasthenics	$c$ (190)	$d$ (310)	$c + d$ (500)
	$a + c$ (350)	$b + d$ (350)	Total (700)

thetical figures). All the statements can now be expressed algebraically. Jung's statement (1) is  $a > c$ . Jung's statement (2) is  $d > b$ . Eysenck's statement, as interpreted above, is

$$d/(c+d) > b/(a+b).$$

Multiplying both sides by  $(a+b)(c+d)$ , this becomes  $da > bc$ . But this is what we get by multiplying together the two relations expressing Jung's two statements. It follows that this version of Eysenck's statement is a correct deduction from Jung's hypotheses. The same deduction can, however, be made from other hypotheses. If Eysenck's statement is true then at least two of the four statements  $a > c$ ,  $d > b$ ,  $a > b$ ,  $d > c$  must be true but we do not know which. Using Foulds's figures  $da > bc$  (Eysenck's statement confirmed), and  $d > b$  (Jung's statement 2 confirmed), but  $a < c$  (Jung's statement 1 disproved).

An alternative formulation of this version of Eysenck's statement,  $da > bc$ , is 'there is a positive association between introversion and dysthymia'. This will be referred to as *the association hypothesis*. If it is true, the four comparisons, such as  $a$  with  $c$ , may be manipulated at will, retaining the association ( $da > bc$ ), by an arbitrary adjustment of the column and row totals. We are not concerned with the relative proportions of hysterics, dysthymics, extraverts and introverts in the population. These are determined by the method of diagnosis and personality assessment used, and by the composition of the population under investigation. It follows that the only statement that is worth testing in this instance is the association hypothesis. If there is an association, it is immaterial which of the four comparisons show it.

Let us suppose that we have a questionnaire which is thought to discriminate between extraverts and introverts, that it is applied to a sample of hysterics and dysthymics, and that a significant association appears in the resulting  $2 \times 2$  table. If we are prepared to assume the validity of the questionnaire, then it is justifiable to conclude that the association hypothesis is confirmed. On the other hand, if we are prepared to assume the truth of the association hypothesis, the evidence is still not quite sufficient to show that the questionnaire is valid. Dysthymia, as Foulds points out, may be associated with variates other than introversion and the questionnaire may involve one of these. Assuming this not to be so, however, one can conclude that the questionnaire is valid. The crux of the argument is that the association hypothesis cannot be proved until we have a valid test of introversion, and vice versa, the validity of a test of introversion cannot be established by this method, until it is shown that the association hypothesis is true.

The argument has so far referred only to an interpretation of Eysenck's statement in terms of proportions. Another interpretation, however, is that dysthymics tend to have higher scores on an introversion scale which is either continuous or has many values. Jung's hypothesis refers to extraverts and introverts, making no mention of degree of introversion. Let us assume that a scale of introversion can be devised and that Jung's extraverts and introverts are defined as those whose scores lie respectively below and above a specified point on the scale. Let the scale and cutting-point be arbitrarily adjusted until the extraverts have a mean score of 5 and the introverts have a mean score of 15. If  $\bar{D}$  and  $\bar{H}$  are the mean scores of dysthymics and hysterics respectively, we see that

$$\bar{D} = \frac{5c + 15d}{c + d}$$

and

$$\bar{H} = \frac{5a + 15b}{a + b}$$

It follows that

$$\bar{D} - \bar{H} = \frac{10(ad - bc)}{(a + b)(c + d)}.$$

We have already seen that it is reasonable to replace Jung's two statements by the association hypothesis, which may be expressed as

$$ad > bc.$$

Thus, if the association hypothesis is true, it follows that

$$\bar{D} - \bar{H} > 0$$

Eysenck's statement, interpreted quantitatively, is therefore a correct deduction from the association hypothesis. The other difficulty remains, however. It is impossible to establish the validity of a test of introversion, using data of this kind, unless the truth of the association hypothesis is assumed.

As far as M.P.I. E-scores are concerned, the statement  $\bar{D} - \bar{H} > 0$  appears to be strongly supported by the evidence available. This is an interesting fact in the comparative study of different categories of neurotic patients but it does not establish that the M.P.I. is a valid index of introversion/extraversion.



of specific receptors. There are also valuable chapters by Livingston on recent developments in neurology which may be of significance in psychology; and by Diamond and Chow on lesions of the receptor areas of the cortex in relation to discrimination, and of the temporal lobe in relation to learning. Ratliff provides an extensive account of the physical and physiological concomitants of vision. But again it is limited on the psychological side to psychophysical data. It includes a comparison of the human dark adaptation curve with that obtained by Blough using reinforcement procedure with pigeons. Many years ago Craik and Vernon found that perception in other than simple discrimination in dim light was determined by other factors as much as by the absolute threshold; one wonders if this is true also for pigeons!

Thus what we really need are more psychological studies to bridge the above-mentioned gap between the data of psychophysical and other perceptual situations. But in this respect there is one really valuable chapter by Attneave, in which, after noting that Graham's S-R approach, though valid enough within its limits, is unlikely to advance the study of perception very much, he proceeds to discuss some other physically measurable variables which might more profitably be studied. He instances particularly the multi-dimensional treatment of a number of variables arising from the phenomenal aspects of percepts; and the significance of classification in establishing perceptual hypotheses. This chapter is useful not only in suggesting profitable new lines of experiment, but also in providing a corrective to the views of Zener and Gaffran on the one hand, and of Ittelson on the other. The former over-estimate the importance of certain phenomenal aspects of perception. It is of course possible that a more systematic study of these might provide effects of other cognitive factors, as the authors tend to do. After all, the main function of perception is to inform us as to what our surroundings are, not merely what they look like. Ittelson and the Transactionalists appear, however, to go to another extreme in considering only perceptual cures. But in relation to individual differences in perception, it is a pity that no study was included of the work of Gardner and his collaborators at the Menninger Clinic on the various dimensions of attention.

It may be gathered, therefore, that in the opinion of the reviewer this book is something of a curate's egg, although it may be admitted that different curates might like different parts of the egg! Some readers might prefer a collection of mainly factual data, such as Rosenblith's *Sensory Communication*. Others may welcome this exposition of different points of view. Still others may like to work out for themselves what it all adds up to, and evolve their own comprehensive theory. But they must be warned that this is likely to be an arduous task.

M. D. VERNON

*The Psychology of Perception.* By M. D. VERNON. Harmondsworth: Penguin Books. 1962. Pp. 265. 5s.

There are many difficulties in writing a popular account of an advancing field of study. One is that of deciding whether to aim at liveliness by introducing the latest material, even though this may be controversial or difficult to integrate with previous views, or on the other hand to try to represent in reasonable perspective a mass of earlier findings which possess at least a semblance of coherence. Prof. Vernon, in this book, has chosen the latter alternative and the outcome is, as one might expect, a remarkably able and balanced survey of her subject. It will serve as a useful (perhaps a trifle too useful?) text for elementary students. It will also be read with profit by a great variety of people who have a serious interest in perception for one reason or another, but lack the technical equipment to tackle more academic books such as the author's own two major works.

The ground covered ranges from the development of perception in children through such topics as perception of objects, shape, colour, space and movement, to more concrete questions about particular types of material of practical import such as pictures, diagrams and print. Later chapters deal with attention and perception, and the relations to perception of motivation and personality. Over 300 notes give copious references to the literature. There are some 37 figures in the text, though no photographs. Considering that the book is exclusively about *visual* perception it is a pity that it could not have been more generously illustrated. As it is, quite a lot of the 240 pages of text are taken up in verbal descriptions of people's visual experience, chiefly in an experimental context. This inevitably makes a considerable extra demand on the

reader's attention, especially as Prof. Vernon quite properly does not resort, as critics and historians of art commonly do, to verbal artifice or turn of phrase to assist the reader's imagination. No one could expect a *Pelican* to be as full of pictures as say Gibson's *The Visual World*, Gombrich's *Art and Illusion* or Metzger's *Gesetze des Sehens*. But seeing the splendid pictorial enterprise shown by *Penguin Books* in many of their recent publications one may feel some legitimate disappointment in this instance.

Throughout the book Prof. Vern sticks with some determination to what might be called the traditional psychological aspects of her subject. Here and there she sketches in briefly and simply some of the physiological factors in such connexions as visual acuity, and image-formation, and the micromovements of the eye. It is not easy to decide whether, however impeccable in themselves, these passages may not generate some misapprehension in the minds of the lay reader. They might suggest that the problems of perception may be divided into two classes, those concerned with events in the eye and the optic pathways on the one hand, and those of the higher, psychological, processes on the other. Yet most of the major preoccupations of contemporary students of visual perception are concerned not with separating these two classes of problem but with re-examining their possible interrelationships. The outcome has often been a realization of the inadequacy of older views about the functions of the more 'peripheral' mechanisms, and a diminished faith in the idea that they can be taken for granted by the psychologist.

However, on the whole, no reader would have any excuse for supposing that he had been presented with a comprehensive and fundamental account of perceptual function. What he has been given is an extremely valuable summary of a wide and coherent, if circumscribed, empirical approach unencumbered with theoretical pre-occupation. Nobody but Prof. Vernon could have given us this and maintained throughout so sustained a balance. In this, as in the quantity of solid information it contains, her book has set a quite fresh standard for the *Pelican Psychology Series*.

R. C. OLDFIELD

*Stereotypy of Imagery and Belief as an Ego Defence.* By ROSEMARY GORDON. British Journal of Psychology Monograph Supplement No. 34. Cambridge University Press, 1962. Pp. vii + 96. 25s.

*Conceptions of Perceptual Defence.* By WILLIAM A. BROWN. British Journal of Psychology Monograph Supplement No. 35. Cambridge University Press, 1962. Pp. viii + 107. 30s.

Dr Gordon's monograph is not easy to assess. Although there is an interesting introductory historical account of stereotypes and a consideration of various aspects of stereotypy as part of the individual's equipment, and an analysis of the nature and function of beliefs in general, the main part of the monograph is concerned with investigations which Dr Gordon herself conducted. From a series of tests, some quite ingenious, Dr Gordon obtained such results as the following. Subjects tended to have either 'controlled' or 'autonomous' imagery. Those with autonomous imagery tended to have more stereotyped images than those with controlled images. Controlled imagers showed more control over rate of reversing of a reversible figure. There was a relationship between stereotypy of images and stereotypy of beliefs. A high degree of stereotypy was frequently accompanied by unsolved Oedipus conflicts, aggressiveness and tendency to projection, while a low degree of stereotypy tended to go with conflict between Id and Superego. In general terms, Dr Gordon suggests that there is 'a very close relationship between the nature and character of a person's ideology and his basic personality and its problems'. It is strange, then, to find that *The Authoritarian Personality* and related work is referred to only casually or indirectly, and Prof. Eysenck receives no mention at all.

As far as the experimental part of the monograph is concerned (and this in effect constitutes the main reason for this monograph's existence), the reader is in real difficulty. Virtually no quantitative data are provided. Correlations are referred to, but their size is not given. The *Belief-Stereotypy Test* is given but not the scoring key. Undoubtedly there is a case for not weighing down a monograph with an excess of quantitative details. But such things as the size of the correlations are essential for understanding what is going on. In the absence of such information, it is very difficult to evaluate Dr Gordon's work at all, and the most one can say is that it is in

places suggestive, but that in the absence of more convincing evidence, most of the conclusions must be regarded as no more than this.

Dr Brown provides us with an excellent and concise critical review of his selected field. More than eight pages of references testify to his thoroughness in the pursuit of relevant literature. He gives careful consideration to the ambiguities involved in the terms 'perceptual defence' and 'perceptual sensitivity', and himself defines 'perceptual defence' as referring to 'any systematic relationship found to hold between stimulus emotionality and the ease of recognition of stimuli'. As he himself later says, 'perceptual defence' may not be the happiest term to use in this sense, including as it does both lowering and raising of recognition thresholds; but at any rate, Dr Brown makes his use of the term quite clear. He goes on to point to what he calls the 'dilemma' of perceptual defence—that while there is evidence that the emotionality of stimuli raises recognition thresholds, there is also evidence that it may lower thresholds. Dr Brown resolves this dilemma by postulating a *curvilinear* relationship between emotionality and thresholds, i.e. argues that relatively mild degrees of emotionality may raise thresholds while more intense degrees of emotionality may lower them. This is really the central point of Dr Brown's monograph, and he argues his case well. The present reviewer, who has also been puzzled by Dr Brown's relationship before accepting it unreservedly. Dr Brown further believes that a greater degree of stimulus emotionality is required to bring recognition thresholds to their maximum with extraverts than with introverts, and he reports an experiment which he himself conducted, the results of course been raised before in various ways, but certainly more attention to this aspect is suggested for such a theory, in terms of positive and negative perceptual response strengths. But as he says, this theory is put forward 'not so much as a substantive explanation of perceptual defence phenomena, but as an illustration of the fact that such an explanation [i.e. in terms of behaviour theory] cannot be deemed impossible'. If Dr Brown has not provided us with a finished theory, he has at least provided us with a remarkably illuminating guide. D. GRAHAM

*Darwinism and the Study of Society: A Centenary Symposium.* Edited by MICHAEL BANTON with an Introduction by J. BRONOWSKI. London: Tavistock. 1961. Pp. xx + 191. 21s.

The twelve essays which compose this volume are based upon papers read at a Conference arranged by the Scottish Branch of the British Sociological Association and held at the University of Edinburgh in 1959. Exactly a hundred years had elapsed since the publication of *The Origin of Species*; and the aim of the symposium was to bring together both biologists and sociologists with a view to assessing the importance and the implications of Darwin's theories for the present-day study of sociological problems.

The topics selected by the various contributors differ widely in their range. Of the biological contributions those which will appeal more especially to the psychologists are those by Prof. Waddington on the 'human evolutionary system', Dr Maynard Smith on 'evolution and history', and Prof. Hogben on 'Darwinism and human society in retrospect'. All three are highly critical of the oversimplified notions of Darwinism adopted by some of the earlier sociological enthusiasts, who conceived it as little more than a synonym for 'a competitive struggle between individuals and between groups'. Dr Barnett outlines the modern view. 'Just as parts of biology have to be built on physics and chemistry, so', he contends, 'some aspects of sociology must be based on biology: the interaction of nature and nurture underlies much of social science; hence social scientists should be familiar with the principles of genetics'.

Within their particular spheres the surveys provided by the sociologists are excellent. The most illuminating is the long paper by Prof. Ginsberg dealing with the distinctive features of social evolution as contrasted with organic. His chief conclusion is that 'mental evolution in man differs radically from mental evolution in the animal world owing to the part played by the social interactions which take place between different minds'. Nevertheless, although the bearing of Darwinian theories on the study of society have thus proved to be far more indirect than was at first assumed, 'there is', he maintains, 'hardly a branch of social inquiry which has remained unaffected'. In concluding the whole series Dr Banton makes some attempt at reconciling the



sharp differences that emerged in the course of the Conference, and points out that, in their present-day researches, modern sociologists seem chiefly interested in two main lines of approach: first, 'the quantitative analysis of social characteristics', as in recent inquiries into the distribution of property or the effects of social mobility; and secondly, 'field investigations and the analysis of case material', undertaken to throw light on such current issues as educational opportunity, occupational prestige, and the influence of social norms on levels of industrial production. These, so he argues, are problems which 'an appeal to biological theories will do little to solve'. They are therefore questions which 'sociologists will have to elucidate without the aid of Darwinism'. However, as Dr Bronowski, shrewdly observes, in thus narrowing their outlook, present-day students of the social sciences may eventually discover that they have 'missed an essential thread through their labyrinth'.

CYRIL BURT

*Introduction to Psychology.* By NORMAN L. MUNN. Boston: Houghton Mifflin. 1962. Pp. xiv + 588. \$6.95.

*Psychology—A Scientific Study of Man.* By FILLMORE H. SANFORD. San Francisco: Wadsworth. London: Prentice-Hall International. 1961. Pp. xxii + 569. 42s.

*Psychology—Principles and Applications.* (Third Edition.) By MARIAN EAST MADIGAN. St Louis: The C. V. Mosby Co. London: Henry Kimpton. 1962. Pp. 379. 41s.

Here are yet three more additions to the range of introductory text-books in psychology—a topic which seems to be crying out for treatment by *Which?* If value for money is the criterion one must reluctantly rule out Munn's *Introduction to Psychology*. Offered as an abridgement of *Psychology: the fundamentals of human adjustment* (4th edition) it represents a reduction of only 27% in length and under 7% in price as compared with the parent volume. The main omissions are of 'neurological material', 'physiological detail', and (surprisingly) 'most of the sensory material of experimental psychology'. The book is designed for a 'briefer course', but to this reviewer the idea seems ill-conceived. Surely no student would grudge the extra 55 cents for the omitted material, even if this material is something the instructor considers unnecessary. And surely the instructor could tell his students what to omit?

On the same criterion Sanford's book comes out very well. The lay-out is even more than usually spacious, and if some of the numerous illustrations will be recognized as old friends, many others are new and charming, but in some cases, unfortunately, superfluous. Other distinctive features include an unusually informative Glossary, and an Author Index incorporating fully documented bibliographical references, such as would delight any editor's heart.

The third book listed is not really new, but is included because previous editions have not been received. Although not so stated on the title-page, nor in the Preface, this text is designed primarily for student nurses. To integrate professional orientation in a course in general psychology is a good idea, which has been ingeniously carried out, but constant reference to the nursing situation—and even to nursing administration—may well irritate the unsuspecting 'general' reader. The format of the book is much less opulent than what one has by now come to take for granted. Yet this sober—by comparison one might even say drab—volume is only one shilling cheaper than the Sanford book, and even here one finds, as an illustration to the statement 'We need friends', a photograph of a row of teenagers at a soda-fountain. No reading references are given, on the grounds that they are 'often so little used by students'. Instead there is a comprehensive list of 'Audio-visual aids'. This, too, could be very useful, but one may perhaps be permitted to wonder whether the 18 min. necessary to view a film entitled 'Chimpanzees reared in darkness' would not be more profitably employed in reading.

B. SEMEONOFF

*Personality—an experimental approach.* By ROBERT W. LUNDIN. New York: The Macmillan Co. 1961. Pp. xiv + 450. 40s.

This book is important, timely, exciting, and at times infuriating. The author is a confirmed follower of Skinner's approach, and as one might expect, a book out of personality, by Skinner is not only a hybrid but may, to some, seem a monstrosity. Essentially what Lundin has done is

this. He has taken the familiar approach of Mowrer, Miller and other neo-behaviourists, who would define personality out of existence by identifying it simply with the sum total of learned responses. This procedure cuts many Gordian knots, and certainly facilitates the experimental analysis of behaviour in terms of learning theory. Lundin's book, therefore, consists for the major part of a clear and well-presented account of modern learning theory as seen by a Skinnerian, together with an attempt to use the facts and principles uncovered in an analysis of some of those topics which traditionally make up the subject matter of books on personality. Lundin, it will be seen, is following in the footsteps of Guthrie, although he does not mention the latter in his bibliography.

Lundin's book has strong and weak points which are rather similar to those shown by Guthrie. The applications of learning theory principles to personality development are often striking and suggestive; clinical psychologists and others working in the field of personality can hardly afford to miss this determined attempt to build a bridge. On the other hand, while much of what is said is sensible, reasonable, and sometimes even probable, nevertheless, the link is purely theoretical and at times semantic. What is one to make of a Skinnerian, who in following his master, should shun theorizing like the plague when he argues that depression is due to lack of reinforcement (p. 69)? Who maintains that over-indulged children are poor risks for later life because they have been too regularly reinforced as children (p. 82)? I am not saying that these hypotheses are wrong; it is quite possible that future research will in fact verify them but it can hardly be argued that at the present time they are anything more than speculative hypotheses, and that there is much evidence suggesting other causes, none of which is reviewed by Lundin. How would he account for the differentiation between psychotic and neurotic depression in terms of sedation thresholds for instance? It is misleading to the reader, particularly the inexperienced one, to present such speculations as if they were factually supported, and a lesser degree of selectivity as far as the evidence is concerned might have been desirable.

In so far as the book has a positive message, I think it deserves considerable recommendation and presents a long due introduction of new ideas into the field of personality. Unfortunately the author has also found it necessary to introduce his subject by a critical discussion of other points of view, which is so inept and lacking in knowledge or understanding of these different approaches that it would have been better if it had been omitted altogether. In his enthusiasm for his environmental approach, Lundin appears to go almost as far as Watson did when he declared that he could turn any child into anything you like provided he had control over what we would now call his 'reinforcement schedule'. There is no serious attempt to argue the case, or to discuss the vast amount of evidence available to contradict such a statement; in an author who stresses the *scientific* approach to personality, such subjectivity and disregard of facts is indeed painful. One can agree with the stress he lays on the importance of conditioning and learning, and the relevance of these agencies in personality development, without disregarding investigations stressing the importance of hereditary factors. Other examples of excessively naïve dismissal of alternative ways of looking at the facts occur in almost every one of the first fifty pages of this book; the reader interested in what Lundin has to contribute would be best advised to start reading at the beginning of Chapter 3. With this proviso I think the book can safely be recommended to psychologists interested in the application of learning theory to personality research.

H. J. EYSENCK

*Determinants of Infant Behaviour.* Edited by B. M. Foss. London: Methuen. 1961.  
Pp. xv + 308. 42s.

This book is a record of the proceedings of a 1959 Tavistock Study Group on Mother-Infant Interaction, and is concerned with this particular topic, rather than with infant behaviour in general. In fact, the main work on the infant-mother relation is in the animal field, the human studies being concerned (except in the first section) rather with maternal deprivation and with institution babies. Some of the contributions have been enlarged and revised for the purposes of publication.

It is not possible here to discuss each paper in detail. Some, such as part of Harlow's work with monkeys, have already (or since) been recorded elsewhere and have indeed found their way into the current editions of psychological text-books, and into television programmes. Others, such as Rheingold's, are a continuation of previously published work.

Part I of this book, titled 'Neonatal Behaviour', includes papers by Gunther, and by Blauvelt and McKenna, which illustrate the importance of simple sucking and orienting behaviour in the building up of the mother infant relationship. The latter paper demonstrates the use of very precise methods of film analysis. This section is perhaps of more interest to the paediatrician than to the psychologist.

The second section includes two papers on animal experiments. One, by Rosenblatt, Turkewitz and Schneirla on 'Early Socialization in the Domestic Cat as based on Feeding and Other Relationships between Female and Young', describes a study of the effects of isolation from the mother on subsequent suckling and socialization. The kittens' responses were found to depend on age at isolation, and on the length of the separation period. The importance of defining the exact circumstances and nature of deprivation in understanding its effects on later behaviour is also brought out in other sections of this book.

The second paper on animal experiments is Harlow's study of 'The Development of Affectional Patterns in Infant Monkeys', and includes some new findings in addition to his well-known comparison of the relative importance of 'clinging' as opposed to 'feeding' in the determination of mother infant relationships. Some of the more recent findings are concerned with 'critical periods', and seem more applicable to human subjects than are his earlier results, which cannot be expected to transfer to subjects who spend more of their time in prams or cradles than in clinging to their mothers in the treetops. Two examples of these later results are, firstly, that monkeys brought up for as few as 60 days without mother or mother-surrogate failed to develop a full attachment and feeling of security from the 'cloth' and 'wire' mothers which had been used as surrogates successfully in other monkeys; secondly, that a 'critical period' for forming attachments to other young monkeys also appeared to exist, for infant monkeys brought up with mother-surrogates only tended to become fixated on them, and to fail in making normal contacts with other infants when finally introduced to them. These findings may be compared with clinical studies of similar situations in human infants.

The third section, the largest, titled 'Social Behaviour', concerns the human infant. It includes case notes by Appell and David on 'Monique', a two-year-old-girl suffering from 'maternal deprivation', who shows very serious mental retardation, thought to be psychogenic rather than organic in origin. The interesting part of this report lies in the account of her rehabilitation, which demonstrates the reversible nature of the symptoms, and gives grounds for some optimism in the treatment of similar cases, if taken in hand sufficiently early.

The same authors report a preliminary study on the effects of 'routine' as opposed to 'intensive individual nursing care' in infants under one year old in a nursery caring for infants undergoing B.C.G. vaccination, separated from their families which have tuberculous infections. This shows a careful consideration of the exact conditions of care in the two groups, with a view to distinguishing between, for example, the effects of isolation (poverty of external stimuli) and those of lack of communication and interaction, often confused in previous studies of institutionalized infants.

Closely related to this paper is one by Rheingold—a continuation of her earlier work on social responsiveness of infants at six months of age. Here she reports work with three-month-old infants; at this stage she finds greater social responsiveness to strangers, as shown by smiling, etc., in institution babies than in other children of this age. The discussion suggests several possible reasons for this finding. Ambrose's paper, which follows, deals with 'The Development of the Smiling Response in Early Infancy' and provides one possible explanation, in that the peak age for smiling at a stranger appears in institution babies seven weeks later than in home babies, indicating that the home infant learns more quickly to distinguish strangers from familiar faces.

The concluding section of the book, consisting of Gewirtz's paper (and a reply to it by Bowlby) makes much more difficult reading than any of the earlier sections, especially for one who is unaccustomed to the terminology of Behaviour Theory. In addition to this 'jargon', Gewirtz tends to use unnecessarily verbose words and phrases such as 'caretaker-interactor' (mother-figure), 'locomotes' (goes), and 'behaviours' (actions). One must admit, however, that the approach of Behaviour theorists does counteract the tendency of some clinical and child psychologists (in the past, especially) to be inexact and 'unscientific'; it is a pity that their writings have to be so indigestible. Some of the members of the study group felt that this attempt to fit the facts into a general theoretical framework was a little premature. Bowlby, however, in his concluding comments on Gewirtz's paper, welcomes its careful definition of the differences



between 'privation' and 'deprivation' as well as its detailed examination of the effects of learning on the growth of attachments, although he himself would place a little more emphasis on the part played by instinct.

In conclusion we can agree that this book provides a useful account of current research methods and results in the field of mother-infant interaction, demonstrating the scientific approach which has replaced the rather haphazard collection of data (often second-hand or retrospective) of former years. One of the chief facts that emerges is the recognition of the need for precise definition of terms and of conditions of research, the lack of which, in earlier studies of 'maternal deprivation', helped to produce apparently confused and conflicting results. A desire for reconciliation of opposing views by members of different theoretical backgrounds is also apparent, in the exchange of ideas in discussion. Although research is still only on the fringe of the problem, and often, for practical reasons confined to animal experiments, progress is gradually being made in understanding the important determinants of infant behaviour.

A. E. NAUGHTON

*Experimental Foundations of Clinical Psychology.* Ed. ARTHUR J. BACHRACH. New York: Basic Books Inc. 1962. Pp. xii + 641. \$10.00.

This symposium of eighteen chapters, most of which are clearly written by psychologists who have a record of good research in their subjects, covers two main areas.

One of them, taking up eight chapters, is concerned with conditioning and learning. It contains clear descriptions of Skinnerian methods of producing desired responses, and of Wolpean methods of eliminating undesired responses. We are given an account of the application of these and other learning procedures to the investigation and modification of various functions: verbal behaviour, psycho-physiological function, infantile experience, perception, and psychotic and neurotic behaviour.

The second main area, taking up six chapters, is concerned with inter-related topics of central importance to clinical psychologists: the measurement of psychological functions, the assessment of psychological disorders within the clinical setting, the interview, psychotherapeutic and group processes, and the modification of beliefs.

The book gives references to detailed, though unvalidated, suggestions on the treatment of neurotic disorders and tension states, and on the modification of certain aspects of psychotic behaviour. Otherwise there is little of immediate clinical relevance.

The main contribution of the book lies in the description of many findings and ideas which are of interest to both research and practical workers. One example is the report that monkeys, operantly conditioned to avoid electric shocks for 6 hr at a time, eventually died of stomach ulcers, while control monkeys, passively experiencing exactly the same shocks, but with no means of avoiding them, remained well. Another example is the suggested application of psycho-physical methods and findings to the investigation of clinical judgement.

The chief limitation of this book lies in the fact that the contributors, like most experimental psychologists, have considered the clinical field largely with the intention of applying their own particular methods and theories. Quite a different approach would seem to be necessary if the 'experimental foundations of psychology' are eventually to be well and truly laid. Such an approach would have to be motivated by a desire to explain, and to treat successfully, psychological disorders which have become thoroughly well known as the result of systematic controlled and uncontrolled observations over some years. In addition, there would have to be a rigorous avoidance of biases which might exclude important observations from consideration. For example, when considering psychotic disorder, one would take into account, among many other things, Pavlov's observations of 'equalization', 'paradoxical', and 'ultra-paradoxical' phases in experimentally induced animal disorder. Likewise, when considering human neuroses, one would not neglect the only substantial source of data: the uncontrolled observations of the psychoanalysts. Again, when examining disorders of perception one would not fail to take into account the findings of Gestalt Psychology.

Such an approach should result in a systematic statement of working hypotheses and of the present limitations of experimental psychology. The achievement of such aims appears, at present, to be beyond our reach. However, their attainment may well be brought nearer if clinical psychologists make a critical study of the book under review.

M. B. SHAPIRO

*Family Constellation: Theory and Practice of a Psychological Game.* By WALTER TOMAN. New York: Springer. 1961. Pp. viii + 248. \$4.50.

The 'Theory' underlying this book is, more or less, that personal relationships in later life are largely determined by sibling patterns. What the 'Game' consists of is never made quite plain, but the implication seems to be that one should approach one's 'family constellation', or those of one's friends, in a spirit of playful inquiry rather than treat the technique as a serious diagnostic aid. Nevertheless, a fairly elaborate 'algebra' and series of 'formulas of family constellation' are offered, in the concluding chapters, as a means of systematizing one's data. Even without having recourse to these quantitative devices, one is struck by the convincingness of the many detailed analyses contained in the main portion of the book, and one is left wondering whether the author has underplayed or overplayed his hand. One possible weakness is that insufficient attention is paid to cases where a predicted outcome has not in fact occurred, i.e. more consideration might have been given to individual deviations within a given pattern. In general, however, this is a stimulating as well as an entertaining little book.

B. SEMEONOFF

*The Explanation of Criminality.* By GORDON TRASLER. London: Routledge and Kegan Paul. 1962. Pp. ix + 134. 20s.

The most important part of this book is its attempt to demonstrate, coherently, comprehensively and with experimental evidence, how the development of some forms of delinquent behaviour can be accounted for in terms of learning theory without recourse to psychopathology, anomie, subcultures, culture-conflicts and so forth. It is not my impression, however, that Dr Trasler is arguing that any of these other concepts can be dispensed with altogether; and it is probably not his fault that the book bears so ambitious a title, rather than 'A possible account of the development of non-pathological forms of delinquency in a British-type society'. I can see why a publisher would prefer the present title.

The argument of the main section of the book can be summarized as follows:

- (1) the learning of non-criminal forms of behaviour involves a process of passive avoidance conditioning, the main conditioned reaction being anxiety;
- (2) the effectiveness of this process depends partly on the strength of this anxiety;
- (3) where there is a strong dependent relationship between child and parents, the sanction or withdrawal of approval evokes intense anxiety;
- (4) the relationship is likely to be one of dependence if it is 'exclusive, affectionate and reliable';
- (5) the effectiveness of the conditioning process also depends on the consistency with which the sanction is applied;
- (6) and on the extent to which it is 'presented in terms of a few well-defined principles';
- (7) extraverts are harder to condition than introverts.

Social class V is over-represented among offenders because its families seldom exemplify 3, 4, 5 or 6. Individual differences in susceptibility to similar degrees of social conditioning are explained in terms of genetic differences in the strength of the reactions of the autonomic nervous system or endocrine system; and Dr Trasler draws attention to studies which suggest that position on the introversion-extraversion continuum is at least partly governed by genetic inheritance. He 'predicts', and with a slight effort produces confirmation, that prisoners are on average more extraverted than non-criminals, although he admits that earlier studies of smaller samples failed to confirm this.

A purist might criticize the tendency to deduce from hypotheses 'predictions' which one knows to be true and to use these as 'confirmations' of the hypothesis. There is, however, at least one place in which Dr Trasler genuinely predicts an observation that has not yet been confirmed or disproved—that because people are not conditioned against motoring offences in childhood it will be found that the rates of motoring offences are not greater in social class V than in other classes.

An exposition of this kind has been needed for a long time, not only so that its assumptions can be clearly understood and examined, but also as a corrective to wholly sociological or psychopathological explanations. From both points of view Dr Trasler's exposition is a good one.

It is true that he might have strengthened some of his evidence; for example, Lykken's observation that a controlled sample of sociopaths showed both less 'anxiety' on a questionnaire and less ability to develop avoidance of 'punished' responses\* would have supported an important step in his argument.

It is a reviewer's role, however, to think up comments which suggest that he knew it all and could have said it better. This reviewer didn't and couldn't, and regards this book as a good buy.

NIGEL WALKER

*On the Threshold of Delinquency.* By JOHN BARRON MAYS, with a preface by T. S. SIMEY. Liverpool: Liverpool University Press. 1959. Pp. xi + 243.

It is rare to find in publications in social science—so often obscured by irrelevant data and side-issues—a clear statement of purpose. In the book under review the author has commendably made such a statement, which is lucidly summarized by Prof. Simey in the preface: 'Mr Mays has sought to discover how far anti-social or emotionally disturbed behaviour is a family and neighbourhood affair, derived from inadequacies in social life which can be remedied by informal education of the families concerned, and the children in them. He has developed a method of achieving this through the combination of conventional techniques of casework and groupwork, which complement rather than supplant the work of the psychotherapist, whose services are still required to deal with the troubles of a minority of deeply disturbed children' (p. viii). Unfortunately, it is equally clear that the stated purpose was not achieved.

This negative outcome is not accidental: the purpose could not be achieved. In Mr Mays's study there was no control group, and no statistical comparison was therefore possible. And although the difference between the principles of psychotherapy and the principles of social work approach are described in some detail, the book contains no indication of the way in which these principles were followed in the boy's club, organized as a 'social therapy group' in a delinquency area. The 'casework-groupwork hypothesis' is stated clearly enough; but there is no evidence in the scientific sense that it was supported or disproved. Also, it is impossible to assess any social work technique when, as Mr Mays admits was the case, the social workers are trying to use the combination of groupwork and casework without having proper training in either (pp. 22 and 25).

Mr Mays does not, however, minimize his shortcoming or his failures. His account is honest and he is not prone to overstating his case. One of the most positive features of the book is a detailed account of the difficulties of ensuring any positive response—let alone cooperation—of the parents whose children the club was serving, and of the children themselves. It appears that the delinquent groups with which Mr Mays was concerned lacked any team spirit necessary for group effort: their games, sport, or self-government were quickly reduced to the level of absurdity. When an attempt was made to institute elected 'house captains' to ensure more cooperation and better discipline, the smallest and youngest boys were made 'captains' so that they could exercise no control over their houses and inter-house games.

Thus the results are meagre. Even where some modest positive findings are reported, they are of doubtful value. Mr Mays states that 'the club seemed to be least successful with those boys who came into membership after having already committed one or more offences. There were twenty-seven such cases of whom ultimately twenty were lost, and only six retained in successful association' (p. 195). But it is apparent from the table on the following page that three out of the six boys subsequently appeared in court on a further charge; and it transpires from the description of various incidents in the club's history that more often than not the court appearance rather than further crime was avoided. Even if we assume that some of the club boys succeeded, it does not follow that 'the club was successful': the boys might have improved despite, rather than because of, club activities.

The above criticisms are methodological. In other respects *On the Threshold of Delinquency* deserves nothing but praise. It is lucidly and vividly written and convincing, the more so because the integrity of the reporting is of the highest order. Events in the club, general observations on aimless destructiveness of the children (and often of their parents), on the short duration of the close relationship between the mother and her children, on the marginal role of the father, and

\* *Studies in Behavior Pathology*, ed. T. R. Sarbin, 1961.



on the social and psychological processes contributing to delinquency, shows acute powers of observation, sympathetic insight and balanced judgement. Such assets are quite enough to offset the weakness of methodology. Mr Mays has made an important contribution to the study of juvenile delinquency.

T. GRYGIER

*A Cure of Delinquents: The Treatment of Maladjustment.* By ROBERT W. SHIELDS. London: Heinemann, 1962. Pp. 191. 21s.

*A Cure of Delinquents* describes the work of a psychotherapist in a residential school for mal-adjusted and delinquent boys over a period of eleven years. This account is contained in seventeen brief chapters laced with occasional excerpts from case histories, and with occasional digressions into such topics as aetiology. This topic which must be thought a rather large subject is handled succinctly in a little over five pages.

It would not be possible to read through the pages of this book without a feeling of growing admiration for its author. Whilst his literary style is not impressive, tending at times to be discursive and repetitive, his very survival is nothing short of astonishing. As the story unfolds, we learn that the therapist has been kicked, punched, scratched, bitten, licked, has had a knife held at his throat, chairs thrown at him, has been threatened with bricks, iron bars, scissors and home-made stilettos, and with attempted strangulation. Nothing of this seems to have discouraged the author either from carrying on with his work or subsequently from writing this book.

The title of the book might suggest that at long last a solution has been found to a difficult social problem. The solution here consists of long-term analytically orientated psychotherapy. Whilst adherents of this method no doubt believe that it cures almost everything, including delinquency, others may be more sceptical. Certainly this book does not give any convincing demonstration, for in an apologetic postscript the author explains that, owing to the closing down of the school, longitudinal follow up study 'which might have lent weight to the argument' was not possible.

In spite of the claim implicit in its title, most readers may find that this book has added little or nothing to their understanding of the causes of delinquency or the treatment of delinquent boys.

E. L. R. MACPHERSON

*The Skills of Interviewing.* By ELIZABETH SIDNEY and MARGARET BROWN. London: Tavistock, 1961. Pp. xii + 396. 35s.

The authors of this book are well qualified, by experience and training, to do the job. They write well and they are obviously very nice people. Much of their experience has been collaborative and it seems proper that this book should be a joint effort.

At first reading the first two chapters seemed unnecessary except as a warming-up exercise for the authors: their twenty-odd pages on the Development of Personality struck me as being thin and inadequate, hustling the reader through a very condensed resumé of the literature, much of it social-anthropological. However, on revision it occurred to me that for the readers they have mainly in mind, 'people engaged in management', this introduction is valuable even if it only referred them to other books.

After page 40 every page is worth reading by all those engaged in interviewing, whether the field is industry and commerce, educational psychology or the practice of psychiatry and clinical psychology.

The highlights for me were, first, the chapter on the basic preparation which the boss must do before he plans to interview applicants. The preparation advised goes away back beyond even the job-description stage: he is reminded, with numerous vivid and convincing examples, that he has to decide whether the job should be done at all by anybody. Secondly, the material used in the book includes transcripts of actual interviews which serve as raw material for object lessons and critical evaluations. Thirdly, the authors steer a course between the bias towards The Management (an American tendency) and one towards The Workers (a British failing.) However, they do not dodge the moral dilemma of the professional person hired by business men and in a chapter on 'Social Responsibilities of the Interviewer' they talk mature good sense—which means I agreed with them.

DENIS MCMAHON

- Le Développement Perceptif.* By ROBERT FRANCÈS. Paris: Presses Universitaires de France. 1962. Pp. iv + 279. NF 16 + T.L.
- La Caractérologie Ethnique.* By PAUL GRIÈGER. Paris: Presses Universitaires de France. 1961. Pp. xx + 319. NF 18 + T.L.
- La Condition de la Femme.* By J. E. HAVEL. Paris: Librairie Armand Colin. 1961. Pp. 224. 4,50 NF.
- Une Recherche sur le Civismisme des Jeunes à la Fin de la Quatrième République.* By JEAN-WILLIAM LAPIERRE and GEORGES NOIZET. Aix-en-Provence: Publication des Annales de la Faculté des Lettres. 1961. Pp. 176. 17 NF.
- Kleine Charakterkunde.* By HUBERT ROHRACHER. Wien-Innesbruck: Urban and Schwarzenberg. 1961. Pp. viii + 299. DM 14; öst.S. 89.
- Guildhall Lectures, 1961.* British Association/Granada. By Sir JAMES GRAY, Professor HERMANN BONDI and Sir JOHN WOLFENDEN. London: University of London Press. 1962. Pp. 80. 4s. 6d.
- The Unicorn Review, Volume 1, Number 1.* Edited by HARRY BELL. Edinburgh and London: W. and R. Chambers Ltd. 1962. Pp. xvii + 80. 5s.
- Rhythmic Activity in Animal Physiology and Behaviour.* Vol. 1. Theoretical and Experimental Biology. By J. L. CLOUDSLEY-THOMPSON. New York and London: Academic Press. 1961. Pp. x + 236. 54s. 6d.
- Kant's Theory of Knowledge. An Outline of One Central Argument in the 'Critique of Pure Reason'.* By GRAHAM BIRD. London: Routledge and Kegan Paul. 1962. Pp. x + 210. 28s.
- The Origins of Science. An Inquiry into the Foundations of Western Thought.* By ERNEST H. HUTTEN. London: George Allen and Unwin. 1962. Pp. 241. 28s.
- Virgin Wives. A Study of Unconsummated Marriages.* By LEONARD J. FRIEDMAN. London: Tavistock Publications. 1962. Pp. xiv + 161. 21s.
- A Dictionary of Symbols.* By J. E. GIRLOT. London: Routledge and Kegan Paul. 1962. Pp. liv + 400. 50s.
- Formal Organizations.* By PETER M. BLAU and W. RICHARD SCOTT. San Francisco: Chandler Publishing Co. 1962. Pp. xii + 312. \$5.25.
- Black Ship to Hell.* By BRIGID BROPHY. London: Seeker and Warburg. 1962. Pp. 490. 50s.
- The Death Doctors.* By A. MITSCHERLICH and F. MIELKE. London: Elek Books. 1962. Pp. 367. 50s.
- The Courage of his Convictions.* By TONY PARKER and ROBERT ALLERTON. London: Hutchinson. 1962. Pp. 192. 16s.
- Hidden Channels of the Mind.* By LOUISA E. RHINE. London: Victor Gollancz. 1962. Pp. xii + 291. 25s.
- Toehold on Zen.* By JEFFREY SWANN. London: George Allen and Unwin. 1962. Pp. 103. 16s.
- Hara—The Vital Centre of Man.* By KARLFRIED GRAF VON DÜRKHEIM. London: George Allen and Unwin. 1962. Pp. 208. 21s.
- The Philosophy of Form.* By GJERLOV-KNUDSEN. Copenhagen: G. E. C. Gad Publishers. 1962. Pp. 169. No price given.
- Le Test des Métamorphoses.* By JACQUELINE ROYER. Lyons and Paris: Emmanuel Vitte. 1961. Pp. 397. No price given.
- Love or Perish.* By SMILEY BLANTON. Kingswood, Surrey: The World's Work (1913) Ltd. 1957. Cedar Special 1962. Pp. 256. 5s.
- The Achieving Society.* By DAVID C. MCCLELLAND. London: Van Nostrand. 1961. Pp. xvi + 512. 60s.
- Corporate Society and Education. The Philosophy of Elijah Jordan.* By GEORGE BARNETT and JACK OTIS. Ann Arbor: The University of Michigan Press. London: The Cresset Press. 1961. Pp. x + 297. 50s.

- Odyssey of the Self-Centred Self.* By ROBERT ELLIOT FITCH. London: George Allen and Unwin. 1962. Pp. viii + 184. 16s.
- Escape from Loneliness.* By PAUL TOURNIER. London: SCM Press. 1962. Pp. 192. 21s.
- The American High School and the Talented Student.* By FRANK O. COPLEY. Ann Arbor: The University of Michigan Press. London: The Cresset Press. 1961. Pp. xx + 92. 25s.
- The American College—A Psychological and Social Interpretation of the Higher Learning.* Edited by NEVITT SANFORD. New York and London: John Wiley. 1962. Pp. xvi + 1084. 75s.
- Boys in White. Student Culture in Medical School.* By HOWARD S. BECKER *et al.* Chicago and London: The University of Chicago Press. 1961. Pp. xiv + 456. 80s.
- Child Rearing in the Lebanon.* By EDWIN TERRY PROTHRO. Cambridge, Mass.: Harvard University Press. London: Oxford University Press. 1961. (Harvard Middle Eastern Monographs VIII.) Pp. viii + 186. 30s.
- Dialogul La Copii.* (Children's Conversation.) By TATIANA SLAMA-CAZACU. Bucharest: Editura Academiei Republicii Populare Romine. 1961. Pp. 167. Lei—5,65.
- The Healing Gods of Ancient Civilizations.* By WALTHER ADDISON JAYNE. New York: University Books, Inc. 1962. Pp. xxii + 573. \$10.00.
- Pavlovian Conference on Higher Nervous Activity.* Edited by NATHAN S. KLINE. Annals of the New York Academy of Sciences. Vol. 92, Art. 3. New York: The N.Y. Academy of Sciences. 1961. Pp. 813–1198.
- Some Biological Aspects of Schizophrenic Behavior.* Edited by D. V. SIVA SANKAR. Annals of the New York Academy of Sciences. Vol. 96. Art. 1. New York: The N.Y. Academy of Sciences. 1962. Pp. 1–490.
- Fundamentals of Psychology: The Psychology of the Self.* Edited by ERNEST HARMS. Annals of the New York Academy of Sciences. Vol. 96. Art. 3. New York: The N.Y. Academy of Sciences. 1962. Pp. 681–894.
- Actualités Neurophysiologiques. Troisième Série.* By A.-M. MONNIER. Paris: Masson et Cie. 1961. Pp. viii + 372. 60 NF.
- Productive Thinking.* By MAX WERTHEIMER. London: Tavistock. 1961. Pp. xviii + 302. 28s. Contains three chapters planned for inclusion in but omitted from the first edition.
- Brett's History of Psychology.* Edited and Abridged by R. S. PETERS. London: George Allen and Unwin. 1962. Pp. 778. 60s.
- Educational Psychology in the Classroom.* (Second Edition.) By HENRY CLAY LINDGREN. London and New York: John Wiley. 1962. Pp. xviii + 574. 53s.
- Conditioned Reflexes.* By I. P. PAVLOV. New York: Dover Publications. 1961. Pp. xvi + 430. \$2.25.
- Psychopathology & Psychiatry. Selected Works.* By I. P. PAVLOV. Moscow: Foreign Languages Publishing House. London: Central Books. 1962. Pp. 543. 21s.
- The Caseworker's Use of Relationships.* By MARGARET L. FERARD and NOËL K. HUNNYBUN. London: Tavistock. 1962. (Mind & Medicine Monographs No. 7.) Pp. xii + 133. 20s.
- The Origins of Science. An Inquiry into the Foundations of Western Thought.* By ERNEST H. HUTTEN. London: George Allen and Unwin. 1962. Pp. 241. 28s.
- Perceval's Narrative. A patient's account of his psychosis, 1830–32.* Edited by GREGORY BATESON. London: Hogarth. 1962. Pp. xxii + 331. 42s.
- A Kierkegaard Critique.* Edited by HOWARD A. JOHNSON and NIELS THULSTRUP. New York: Harper. 1962. Pp. viii + 311. 42s.
- They Shall Take Up Serpents. Psychology of the Southern Snake-Handling Cult.* By WESTON LA BARRE. Minneapolis: Minnesota University Press. London: Oxford University Press. 1962. Pp. vi + 208. 30s.
- The Etiology of Idiopathic Epilepsy.* By HAROLD GEIST. New York: Exposition Press. 1962. Pp. 297. \$5.00.



- Childhood Schizophrenia.* By WILLIAM GOLDFARB. Cambridge, Mass.: Harvard University Press for Commonwealth Fund. London: Oxford University Press. 1961. Pp. xxvi + 246. 36s.
- Mental Radio.* By UPTON SINCLAIR. (Revised Second Printing.) Springfield, Ill.: Thomas. 1962. Pp. xvi + 237. \$8.50.
- Interpersonal Competence and Organizational Effectiveness.* By CHRIS ARCYRIS. Homewood, Ill.: Irwin. 1962. Pp. xi + 292. \$6.50.
- Some Human Oddities and Very Peculiar People.* By ERIC J. DINGWALL. New York: University Books, Inc. 1962. Pp. 198, 223, respectively. \$6.00 each.
- Epilegomena to the Study of Greek Religion and Themis.* By JANE ELLEN HARRISON. New York: University Books. 1962. Pp. lvi + 600. \$10.00.
- Current Projects in the Prevention, Control and Treatment of Crime and Delinquency.* New York: National Research and Information Center on Crime and Delinquency. National Council on Crime and Delinquency. 1962. Pp. 659. \$6.75.
- Oscillation Experienced in the Perception of Figures.* By KRISTIAN HOLT-HANSEN. Denmark: Ejnar Munksgaard. (*Hist. Filos. Medd. Dan. Vid. Selsk.* 39, No. 7.) 1962. Pp. 48. Kr. 9.00.
- Body, Mind, and the Sensory Gateways.* By FELIX DEUTSCH. New York and Basel (Switzerland): Karger. (Volume 2 of *Advances in Psychosomatic Medicine*.) 1962. Pp. iv + 106. sFr. 18.
- Methoden und Ergebnisse der Psychologischen Unfallforschung.* By ERICH MITTENECKER. Wien: Franz Deuticke. 1962. Pp. viii + 168. öst.S. 151.
- Psychologie de la Motivation.* (Second Edition.) By PAUL DIEL. Paris: Presses Universitaires de France. 1962. Pp. viii + 325. NF 12 + T.L.
- La Pensée Causale.* By MONIQUE LAURENDEAU and ADRIEN PINARD. Paris: Presses Universitaires de France. 1962. Pp. xii + 228. NF 15 + T.L.
- La Notion de Rôle en Psychologie Sociale.* By ANNE-MARIE ROCHEBLAVE-SPENLÉ. Paris: Presses Universitaires de France. 1962. Pp. 434. NF 28 + T.L.
- L'Epistémologie Positive et la Critique Meyersonienne.* By GEORGES MOURÉLOS. Paris: Presses Universitaires de France. 1962. Pp. 230. NF 18 + T.L.



















